

XX Claim 7, Page 64; 70pp; English.
 PS
 CC The present sequence represents a human Tc1-1b protein. The Tc1-1b gene
 CC is implicated in the development of T cell malignancies. Fragments of
 CC Tc1-1b cDNA sequences are used for detecting a target sequence
 CC indicating a chromosome 14 abnormality, such as a (14:14)(q11;q32)
 CC translocation or a (14)(q11;q32) inversion. Tc1-1b antisense sequences
 CC and antibodies are useful for treating a disease state such as T-cell
 CC leukemia or lymphoma associated with a chromosome 14 abnormality. The
 CC Tc1-1b gene and its gene product are useful for treating disease states
 CC associated with the Tc1-1b locus on chromosome protein including
 CC T-prolymphocytic leukaemias, acute and chronic leukaemias associated
 CC with the immunodeficiency syndrome ataxia-telangiectasia (AT).

XX Sequence 128 AA;

Query Match 100.0%; Score 685; DB 21; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1e-72;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MASEASVRLGVPGRMLIQPGIYDESGRTWTVVVRNPSSRRMARASQSGRYEPSIT 60
 Db 1 maseasvrlgvpgrmlwipgylvedeegrtwvvrnpssrrmarasqsgsryepsit 60
 OY 61 VHLQMAMVHRELLSSGMPFSQLPAYWQLYGRKRYRADSFWEIADHGOIDSMEQLVL 120
 Db 61 vhlqmamvhtrellssgmpfsqlpaywqlypgrkryraadsfweiadhgidsmeqlvl 120
 OY 121 TYPERKD 128
 Db 121 typerkd 128

RESULT 2

AAR94976
 ID AAR94976 standard; Protein; 108 AA.

XX AAR94976;

XX 23-AUG-1996 (first entry)

XX Mature T-cell proliferative 1 protein.

XX Tc1-1; chromosome-14; leukaemia; lymphoma; therapy; diagnosis;
 XX T-cell proliferative 1 protein.

XX Homo sapiens.

XX Key Location/Qualifiers
 XX Misc-difference 108 /note="unidentified amino acid"

XX WO9613514-A1.

XX 09-MAY-1996.

XX 23-OCT-1995; 95WO-US13663.

XX 27-OCT-1994; 94US-0330272.

XX (RAGG-) RAGGIO-ITALGENE SPA.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Croce CM, Russo G;

XX WPI; 1996-239444/24.

XX New Tc1-1 protein and gene associated with chromosome 14
 XX abnormalities - useful to develop prods. for detection, treatment
 XX and prevention of diseases such as T-cell leukaemia(s) and
 XX lymphoma(s)

XX Disclosure; Page 69; 105pp; English.

PS
 CC The mature T-cell proliferative 1 protein (AAR94976) is involved
 CC in a translocation of chromosome 14 and X chromosome
 CC t(14:X)(q11;q28). It shows 408 amino acid sequence homology to
 CC human Tc1-1 protein (AAR94974), which is associated with
 CC chromosome 14 abnormalities leading to leukaemia and lymphoma.

XX Sequence 108 AA;

Query Match 25.8%; Score 176.5; DB 17; Length 108;
 Best Local Similarity 33.6%; Pred. No. 4.2e-13;
 Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

OY 1 MASEASVRLGVPGRMLIQPGIYDESGRTWTVV--VRFNPSRRMARASQSGRYEP 57
 Db 1 mage--dvgsapbhlwvbnqgilyrdeyqtrwvaveetsflrarvqldgdaarp 57
 OY 58 SITVHLMQAMVHRELLSSGMPFSQLPAYWQLYGRKRYRADSFWEIADHGOIDSMEQ 117
 Db 58 s-----hlt-----sqldlmwqlypeerymdnstrlwqldhlmvrygqe 98
 OY 118 LVLTYPQPE 125
 Db 99 lllkllpd 106

RESULT 3

AAR94974
 ID AAR94974 standard; Protein; 113 AA.

XX AAR94974;

XX 23-AUG-1996 (first entry)

XX Human Tc1-1 polypeptide.

XX Tc1-1; chromosome-14; leukaemia; lymphoma; therapy; diagnosis.

XX Homo sapiens.

XX Key Location/Qualifiers
 XX Modified-site 5 /label= Phosphorylation_site
 XX /note="casein kinase II phosphorylation site"

XX WO9613514-A1.

XX 09-MAY-1996.

XX 23-OCT-1995; 95WO-US13663.

XX 27-OCT-1994; 94US-0330272.

XX (RAGG-) RAGGIO-ITALGENE SPA.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Croce CM, Russo G;

XX WPI; 1996-239444/24.

XX N-PSDB; AAT18876.

XX New Tc1-1 protein and gene associated with chromosome 14

XX abnormalities - useful to develop prods. for detection, treatment

XX and prevention of diseases such as T-cell leukaemia(s) and

XX lymphoma(s)

PS Claim 6; Page 68; 105pp; English.

XX The amino acid sequence of human Tc1-1 protein (AAR94974) was
 CC deduced from a cDNA clone (AAT18876) obt'd. from an ALL cell line cDNA

DR N-PSDB; AA160765.
 XX Novel nucleic acids and polypeptides, useful for treating disorders
 PT such as central nervous system injuries -
 XX
 XX Example 2: SEQ ID NO 6540; 10078bp; English.
 PS
 CC The invention relates to human nucleic acids (AA157798-AA161369) and
 CC the encoded polypeptides (AA158642-AA162213) with neurotropic,
 CC immunosuppressant and cytostatic activity. The polynucleotides are useful
 CC in gene therapy. A composition containing a polypeptide or polynucleotide
 CC of the invention may be used to treat diseases of the peripheral nervous
 CC system, such as peripheral nervous injuries, peripheral neuropathy and
 CC localised neuropathies and central nervous system diseases, such as
 CC Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
 CC utilisation of the activities such as: immune system suppression,
 CC Activin/Inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,
 CC assays for receptor activity, arthritis and inflammation, leukaemias and
 CC C.N.S disorders.
 CC Note: The sequence data for this patent did not form part of the printed
 CC specification.
 XX
 SQ Sequence 129 AA:
 Query Match 21.8%; Score 149; DB 22; Length 129;
 Best Local Similarity 29.2%; Pred. No. 9, 1e-10;
 Matches 33; Conservative 21; Mismatches 45; Indels 14; Gaps 1;
 QY 13 PGRMTWIOBPGIYDEDEGRFTWTVVVRNPSRRERARASQGSRYESTIVHMQMAVHTGE 72
 Db 30 pdr1awakfkyldkqhapltleikdrq-----lvvllrredvlgrr 75
 QY 73 LLSGGMPFSQLPANWQLYPGRKRYRADSSFWELADHGQIDSEOLVLTQPE 125
 Db 76 pmtcpqigspllpimwqlypdgryrsdsstwrlyvnhkldgvedmllelppd 128
 RESULT 6
 AAR94975 standard; Protein; 113 AA.
 ID AAR94975
 AC AAR94975;
 XX 23-AUG-1996 (first entry)
 DE Human TCL-1 polypeptide.
 XX
 XX TCL-1; chromosome-14; leukaemia; lymphoma; therapy.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FH Misc-difference 31
 FT Misc-difference /note= "unidentified amino acid"
 FT Misc-difference 46
 FT Misc-difference /note= "unidentified amino acid"
 FT Misc-difference 47
 FT Misc-difference /note= "unidentified amino acid"
 FT Misc-difference 55
 FT Misc-difference /note= "unidentified amino acid"
 XX
 XX MO9613514-A1.
 XX
 XX 09-MAY-1996.
 XX
 XX 23-OCT-1995; 95WO-US13663.
 XX
 XX 27-OCT-1994; 94US-0330272.
 XX
 XX (RAGG-) RAGGIO-ITALGENE SPA.
 PA

PA (UYJE-) UNIV JEFFERSON THOMAS.
 XX
 XX Croce CM, Russo G;
 XX WPT; 1996-239444/24.
 DR N-PSDB; AAT18877.
 XX
 XX New TCL-1 protein and gene associated with chromosome 14
 PT abnormalities - useful to develop prods. for detection, treatment
 PT and prevention of diseases such as T-cell leukaemia(s) and
 PT lymphoma(s)
 PS Disclosure; Page 69-72; 105pp; English.
 XX
 XX The amino acid sequence (AAR94975) of human TCL-1 protein was
 CC deduced from the exon sequences of the TCL-1 gene (AAT18877).
 CC A sequence deduced from a cDNA clone is given in AAR94974.
 CC The TCL-1 gene is expressed at high levels in leukaemic cells
 CC carrying a t(14;14)(q11;q32) translocation or an inv(14)(q11;q32)
 CC inversion. The TCL-1 protein, and antibodies raised against it,
 CC can be used for the diagnosis or treatment of conditions associated
 CC with increased expression of TCL-1 proteins, and/or with chromosome
 CC 14 abnormalities, esp. T-cell leukaemia and lymphoma.
 XX
 SQ Sequence 113 AA:
 Query Match 19.3%; Score 132; DB 17; Length 113;
 Best Local Similarity 26.8%; Pred. No. 7, 6e-08;
 Matches 33; Conservative 17; Mismatches 39; Indels 34; Gaps 3;
 QY 13 PGRMTWIOBPGIYDEDEGRFTWTVVVRNPS-----RRE---MARASQGSRYESTIVH 62
 Db 14 pdr1awakfkyldkqhapltleikdrqxxvllrredvlgrrpmtlxgpl--- 70
 QY 63 LMQMAVHTRELLSGMPFSQLPANWQLYPGRKRYRADSSFWELADHGQIDSEOLVLTQ 122
 Db 71 -----lpimwqlypdgryrsdsstwrlyvnhkldgvedmllel 109
 QY 123 QPE 125
 Db 110 lpd 112
 RESULT 7
 AAE05866 standard; Protein; 513 AA.
 ID AAE05866
 AC AAE05866;
 XX 24-SEP-2001 (first entry)
 DE Pseudomonas stutzeri open reading frame-O (ORF-O) protein.
 XX
 XX Pseudomonas stutzeri
 KW Open reading frame-O; ORF-O; pyridine-2,6-bis (thiocarboxylate);
 KW PUTC; environmental remediation; phytoremediation; bioaccumulation;
 KW water purification; solution mining mobilisation; immobilisation;
 KW detoxification; redox state modifier; metal ion reactivity;
 KW carbon tetrachloride; metal.
 XX
 XX Pseudomonas stutzeri.
 OS
 XX
 XX WO200153309-A1.
 XX
 XX 26-JUL-2001.
 XX
 XX 19-JAN-2001; 2001WO-US02386.
 XX
 XX 20-JAN-2000; 2000US-0177251.
 XX
 XX (IDAH-) IDAHO RES FOUND INC.
 XX (PASZ/) PASZCZYNSKI A.
 XX (SEBA/) SEBAT J L.
 PA

PT	branching enzyme sequences	
XX		
PS	disclosure; Fig 8; 108pp; English.	
XX		
CC	The present sequence represents a potato starch branching enzyme (SBE).	
CC	The specification describes a method for affecting enzymatic activity in	
CC	a plant. The method comprises expressing in the plant a nucleotide	
CC	sequence encoding, partially or completely, an intron of a class A potato	
CC	SBE in a sense orientation, optionally together with a nucleotide	
CC	sequence which codes, partially or completely, for an intron of a class B	
CC	SBE in a sense or antisense orientation, and where the nucleotide	
CC	sequence does not contain a sequence that is sense to an exon sequence	
CC	generally associated with the intron. The method can be used for providing	
CC	genetically modified plants which are capable of producing modified and	
CC	improved starches whose properties would satisfy various industrial	
CC	requirements. They can be used for preparing tailor-made starches in	
CC	plants which could replace the post-harvest modified starches. They can	
CC	also be used for expressing genes of interest, e.g. pharmaceutically	
CC	active proteins or enzymes or proteins which are beneficial to plants.	
XX		
SQ	Sequence 906 AA.	
Query Match	10.28; Score 70; DB 19; Length 906.	
Best Local Similarity	24.28; P-val 0.246	

Query Match	10.28;	Score 70;	DB 19;	Length 906;
Best Local Similarity	24.28;	Pred. No. 24;		
Matches 31;	Conservative 17;	Mismatches 34;	Indels 46;	Gaps 8

QY	17	MIQPPGTYDEEGRTWTVTYVVRNPSRERMAA-SQGSRYEPSTVHLMQAVHTRE---	72
Db	621	widp-----regnmw-----sydkcrtypladsenhlytk--fmaafdramsidekf	667
QY	73	-LLSSGMPFS-----QLPAVWQLYPRG-----KYRAA-DSSE	103
Db	668	sflasqgiqvssmdddnkvvfergdlvfnfnphkntyeykvgcdlpgkkyvadsda	727
QY	104	WEIDHGO 111	
Db	728	wefqbar 735	

	RESULT	9
ID	AA69300	
XX	AA69300 standard; Protein; 906 AA.	
AC	AA69300;	
XX		
DT	07-DEC-1998 (first entry)	
XX		
DE	Potato class B starch branching enzyme.	
XX		
KW	Starch branching enzyme; SBE; potato; antisense; amylopectin;	
KW	transgenic plant.	
XX		
OS	Solanum tuberosum cv. Desiree.	
XX		
PN	M09837213-A1.	
XX		
PD	27-AUG-1998.	
XX		
PF	23-FEB-1998; 98MO-IB00270.	
XX		
PR	24-MAR-1997; 97GB-0006060.	
PR	21-FEB-1997; 97GB-0003663.	
XX		
PA	(DANT-) DANISCO AS.	
XX		
PI	Poulsen P;	
XX		
DR	WPI; 1998-467573/40.	
DR	N-PADB; AAV55069.	
XX		

Affecting enzymatic activity using antisense intron inhibition -
especially of starch branching enzyme in plants. useful to alter
PT

PT starch branching enzyme activity, amylopectin levels or starch composition

XX Example: Fig 12; 95pp; English.

CC This is the amino acid sequence of the class B starch branching enzyme (SBE) of potato cv. Desiree. It was deduced from a full-length SBE gene sequence (see AAV55069). A novel method of affecting enzymatic activity in plants or other starch-producing organisms (or their cells, tissues or organs) comprises expressing a nucleotide sequence coding for an intron (see AAV55046) of a class A potato SBE in an antisense orientation, but not containing a sequence antisense to an exon sequence normally associated with the intron. The sequence is optionally expressed with a nucleotide sequence which codes, partially or completely, for an intron of a class B SBE in an antisense/sense orientation. Also claimed are a promoter (see AAV55047) useful in expressing the nucleotide sequence of the method, a transgenic starch-producing organism, and the starch produced by the organism. The method and sequences enable production of transgenic starch-producing organisms with altered characteristics, and particularly plants with altered SBE activity, amylopectin levels and/or starch composition.

CC Sequence 906 AA;

Query Match 10.2%; Score 70; DB 19; Length 906;

Best Local Similarity 24.2%; Pred. No. 24; Matches 31; Conservative 17; Mismatches 34; Indels 46; Gaps 8;

OY 17 WIOPGIYEDGRTWVYVVFNFPSRREMAR-SQGSRYEPSTVHLQMVAHTRE--- 72
 Db 621 Wldfp-----regnw-----sydkorqvwlnadsehlyrk---fmafdrammsldexf 667
 OY 73 -LSSGOMPFS-----QLPAVWQLYPR-----KYRAA-DSSP 103
 Db 668 sfiasgkqivssmdndkvvfergdlvfvfnfphntkygykvgcdipgyrvalsdsd 727
 OY 104 WEIADHGO 111
 Db 728 wefgghyr 735

RESULT 10

AAW14740 ID AAW14740 standard; Protein: 476 AA.

XX AAW14740;

XX 12-MAY-1997 (first entry)

DE Protein encoded by IBV clone GLS-3.

XX Infectious bursal disease virus; IBV; large segment; GLS 5 strain;

KW structural protein; VP2; VP3; VP4; Immunogen; vaccine;

KM Gumboro disease.

OS Infectious bursal disease virus.

PN US5595912-A.

XX 21-JAN-1997.

XX 04-MAY-1990; 90US-0519202.

XX 23-MAR-1994; 94US-0216276.

XX 04-MAY-1990; 90US-0519202.

XX 28-JUN-1993; 93US-0083784.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX Snyder D, Vakharlia V;

XX PI

DR WPI: 1997-107584/10.

DR N-PSDB; AAT63068.

XX Infectious bursal disease virus nucleic acids - for prodn. of

PT vaccines for poultry for protection against Gumboro disease

XX Claim 5; Column 71-76; 41pp; English.

CC The sequences given in AAW14738-41 are encoded by fragments of the infectious bursal disease virus (IBV) large segment from the GLS 5 strain. These sequences are encoded by clones GLS-1 to GLS-4. Proteins such as these, derived from IBV, esp. VP2 which is the major host protective immunogen of IBV, may be used to prepare vaccines against IBV infection/Gumboro disease.

CC Sequence 476 AA;

Query Match 10.1%; Score 69; DB 18; Length 476;

Best Local Similarity 23.8%; Pred. No. 13; Matches 30; Conservative 18; Mismatches 50; Indels 28; Gaps 5;

OY 8 RUGVPRGLMTORPGIYEDGRTWVYVVFNFPSRREMAR-----ASQGSRYEP 57
 Db 162 Rlg-----lkaagpafdcvncgpnwafikrfrhnpwdrlpylnlpylppagrqy-- 214
 OY 58 SITVHLQMVAHTREL--LSSGOMPFSQLPVWQLYPRGKRYRAADSSFWELADHGOIDSM 115
 Db 215 ----lhamasefkehpelasesavrameaasvqpl-----lqsalvfmwleengivldm 265

OY 116 EQLVLT 121
 Db 266 anfaIs 271

RESULT 11

AAW14741 ID AAW14741 standard; Protein: 540 AA.

XX AAW14741;

XX 12-MAY-1997 (first entry)

DE Protein encoded by IBV clone GLS-4.

XX Infectious bursal disease virus; IBV; large segment; GLS 5 strain;

KW structural protein; VP2; VP3; VP4; Immunogen; vaccine;

KM Gumboro disease.

OS Infectious bursal disease virus.

PN US5595912-A.

XX 21-JAN-1997.

XX 04-MAY-1990; 90US-0519202.

XX 23-MAR-1994; 94US-0216276.

XX 04-MAY-1990; 90US-0519202.

XX 28-JUN-1993; 93US-0083784.

XX (UYMA-) UNIV MARYLAND BALTIMORE.

XX Snyder D, Vakharlia V;

XX PI WPI: 1997-107584/10.

XX N-PSDB; AAT63069.

XX Infectious bursal disease virus nucleic acids - for prodn. of

PT vaccines for poultry for protection against Gumboro disease

XX Claim 2; Column 65-70; 41pp; English.

XX PI

The sequences given in AAM14738-41 are encoded by fragments of the infectious bursal disease virus (IBDV) large segment from the GLS 5 strain. These sequences are encoded by clones GLS-1 to GLS-4. Proteins such as these, derived from IBDV, esp. VP2 which is the major host protective immunogen of IBDV, may be used to prepare vaccines against IBDV infection/Gumboro disease.

Sequence 540 AA;

Query Match 10.1%; Score 69; DB 18; Length 540;
Best Local Similarity 23.8%; Pred. No. 16;
Matches 30; Conservative 18; Mismatches 50; Indels 28; Gaps 5;

OY 8 RLGVPGRLMIORGIYDEDEGRVTWTVVFRNPSRREMAR-----ASGGRYEP 57
DB 403 rlg-----lklagpafdvntgpnwatfkrfphnprdwrlpylnlpylppnagry-- 455
OY 58 SITVHLMQMAVHTREL--LSSGMPFSQLPVWQLYGKRYRAADSSFWETADHGQIDSM 115
DB 456 ----hamaasefkepclesavrameaasvdp1-----fgsalvfmwleenglyvldm 506
OY 116 EQLVLT 121
DB 507 anfalls 512

RESULT 12

AAR85439
ID AAR85439 standard; Protein; 1012 AA.

AC AAR85439;

DT 15-FEB-1996 (first entry)

DE IBDV GLS structural protein fragment.

KW IBDV, immunogen; vaccine; poultry.

OS Infectious bursal disease virus of chickens.

PN W09526196-A1.

PD 05-OCT-1995.

PF 29-MAR-1995; 95MO-US03772.

PR 29-MAR-1994; 94US-0219262.

PA (UYMA-) UNIV MARYLAND BALTIMORE.

PI Mengel-Whersat SA, Snyder DB, Vakharina V;

DR WPI; 1995-351197/45.

DR N-PSDB; AAT05623.

PT Chimeric infectious bursal disease virus polypeptide immunogens
comprising VP2 with modified epitopic amino acid sequence for
preparation of vaccines for immunisation of poultry

PS Disclosure; Fig 5A-K; 77pp; English.

CC A CDNA clone contg. the entire coding region of the large RNA
segment of infectious bursal disease virus serotype I strain GLS
has the sequence given in AAT05623, and encodes the structural
proteins VP2/4/3 (AAR85439). Chimeric immunogens in which an
epitopic determinant of a lethal IBDV strain is inserted into
the GLS VP2 can be expressed in recombinant mammalian and poultry
host cells.

Sequence 1012 AA;

Query Match 10.1%; Score 69; DB 16; Length 1012;
Best Local Similarity 23.8%; Pred. No. 36;
Matches 30; Conservative 18; Mismatches 50; Indels 28; Gaps 5;

OY 8 RLGVPGRLMIORGIYDEDEGRVTWTVVFRNPSRREMAR-----ASGGRYEP 57
DB 698 rlg-----lklagpafdvntgpnwatfkrfphnprdwrlpylnlpylppnagry-- 750
OY 58 SITVHLMQMAVHTREL--LSSGMPFSQLPVWQLYGKRYRAADSSFWETADHGQIDSM 115
DB 751 ----hamaasefkepclesavrameaasvdp1-----fgsalvfmwleenglyvldm 801
OY 116 EQLVLT 121
DB 802 anfalls 807

RESULT 13

AAM14737
ID AAM14737 standard; Protein; 1012 AA.

AC AAM14737;

DT 08-MAY-1997 (first entry)

DE IBDV strain GLS large segment derived protein.

KW Infectious bursal disease virus; IBDV, large segment; GLS 5 strain;
structural protein; VP2; VP3; VP4; Immunogen; vaccine;
Gumboro disease.

OS Infectious bursal disease virus.

PN US5595912-A.

PD 21-JAN-1997.

PF 04-MAY-1990; 90US-0519202.

PR 23-MAR-1994; 94US-0216276.

PR 04-MAY-1990; 90US-0519202.

PR 28-JUN-1993; 93US-0083784.

PA (UYMA-) UNIV MARYLAND BALTIMORE.

PI Snyder D, Vakharina V;

DR WPI; 1997-107584/10.

DR N-PSDB; AAT63065.

PT Infectious bursal disease virus nucleic acids - for prodn. of
vaccines for poultry for protection against Gumboro disease

PS Claim 1; Column 43-52; 41pp; English.

CC This sequence is encoded by a fragment of the infectious bursal disease
virus (IBDV) large segment from the GLS 5 strain. This sequence
represents the structural proteins VP2, VP3 and VP4. These proteins
esp. VP2 which is the major host protective immunogen of IBDV, may be
used to prepare vaccines against IBDV infection/Gumboro disease.

Sequence 1012 AA;

Query Match 10.1%; Score 69; DB 18; Length 1012;
Best Local Similarity 23.8%; Pred. No. 36;
Matches 30; Conservative 18; Mismatches 50; Indels 28; Gaps 5;

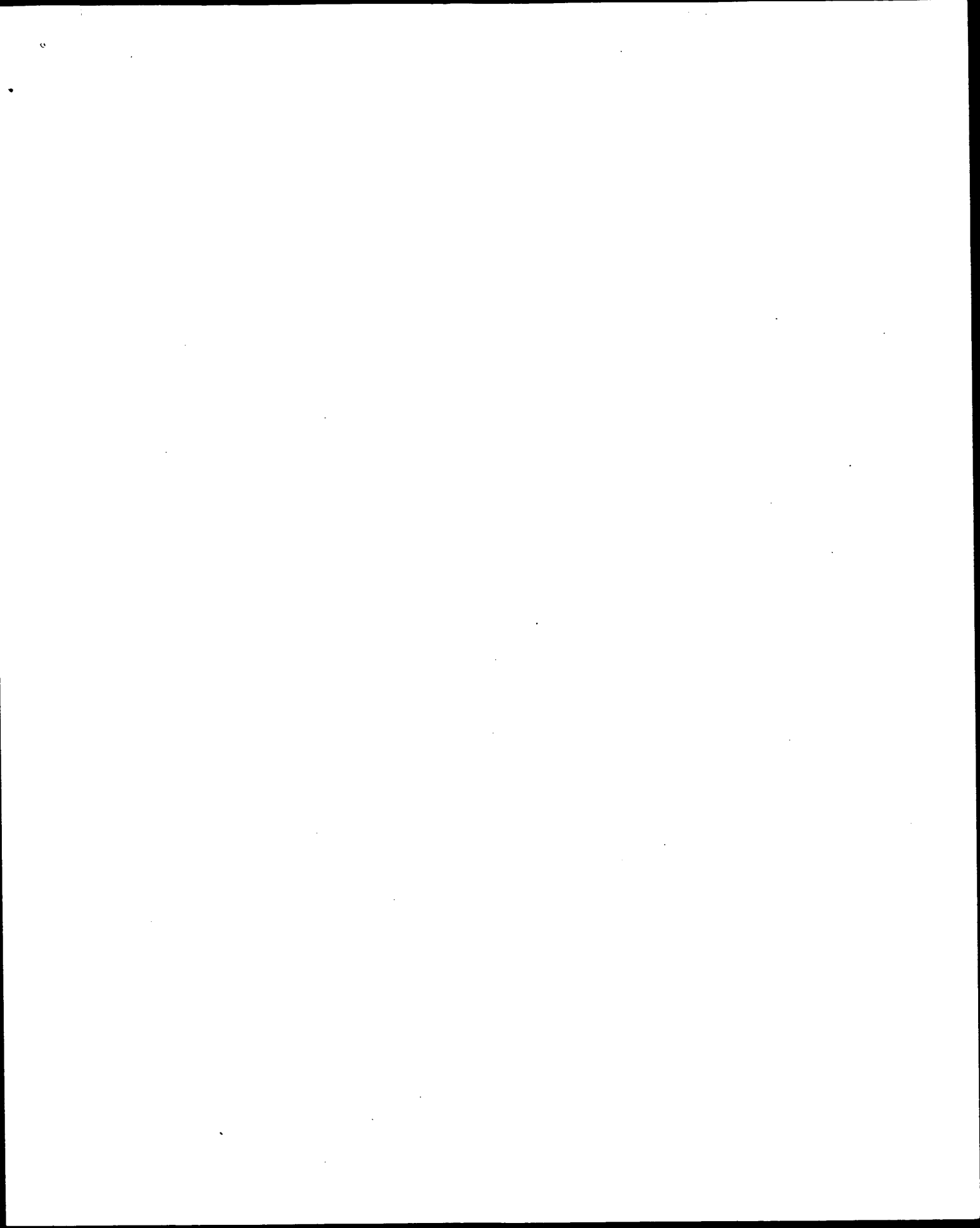
OY 8 RLGVPGRLMIORGIYDEDEGRVTWTVVFRNPSRREMAR-----ASGGRYEP 57
DB 698 rlg-----lklagpafdvntgpnwatfkrfphnprdwrlpylnlpylppnagry-- 750
OY 58 SITVHLMQMAVHTREL--LSSGMPFSQLPVWQLYGKRYRAADSSFWETADHGQIDSM 115

o Fri Nov 30 11:03:09 2001

us-09-526-329-39.rag

Page 9

Search completed: November 29, 2001, 04:03:10
Job time: 516 sec



GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 03:58:04 ; Search time 14.71 Seconds

(without alignments)
195,814 Million cell updates/sec

Title: US-09-526-329-39

Perfect score: 685
Sequence: 1 MASEASVRLGVPGRMTWQR.....HGQIDSMEDVLVTPERKD 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 212252 seqs, 22503292 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents: AI.*
1: /cgn2_6/ptodata/2/1aa/5A.COMB.pep.*
2: /cgn2_6/ptodata/2/1aa/5B.COMB.pep.*
3: /cgn2_6/ptodata/2/1aa/6A.COMB.pep.*
4: /cgn2_6/ptodata/2/1aa/6B.COMB.pep.*
5: /cgn2_6/ptodata/2/1aa/PCMTUS.COMB.pep.*
6: /cgn2_6/ptodata/2/1aa/bockfilest1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	176.5	25.8	108	2	US-08-330-272-4
2	176.5	25.8	108	5	PCT-US95-13663-4
3	149	21.8	113	2	US-08-330-272-2
4	149	21.8	113	5	PCT-US95-13663-2
5	69	10.1	476	1	US-08-216-276A-33
6	69	10.1	476	1	US-08-216-276A-31
7	69	10.1	1012	1	US-08-216-276A-19
8	69	10.1	1012	1	US-08-219-262B-1
9	69	10.1	1012	1	US-08-219-262B-9
10	69	10.1	1012	1	US-08-219-262B-12
11	69	10.1	1012	3	US-09-031-655-1
12	69	10.1	1012	3	US-09-031-655-9
13	69	10.1	1012	3	US-09-031-655-12
14	69	10.1	1012	2	US-08-708-541A-34
15	67.5	9.9	1012	2	US-08-219-262B-7
16	67.5	9.9	1012	4	US-09-031-655-7
17	67.5	9.9	3072	4	US-09-413-814-93
18	67.5	9.9	3079	4	US-09-413-814-80
19	66	9.6	1012	1	US-07-944-525-2
20	66	9.6	1012	1	US-08-219-262B-3
21	66	9.6	1012	1	US-08-219-262B-4
22	66	9.6	1012	1	US-08-219-262B-5
23	66	9.6	1012	1	US-08-219-262B-6
24	66	9.6	1012	1	US-08-219-262B-8
25	66	9.6	1012	2	US-08-708-541A-30
26	66	9.6	1012	3	US-09-031-655-3
27	66	9.6	1012	3	US-09-031-655-4

28	66	9.6	1012	3	US-09-031-655-5	Sequence 5, Appl
29	66	9.6	1012	3	US-09-031-655-6	Sequence 6, Appl
30	66	9.6	1012	3	US-09-031-655-8	Sequence 8, Appl
31	65	9.5	276	2	US-08-712-072C-4	Sequence 4, Appl
32	65	9.5	797	2	US-08-663-566A-2	Sequence 2, Appl
33	65	9.5	797	2	US-08-023-610-2	Sequence 2, Appl
34	65	9.5	797	2	US-08-288-065A-2	Sequence 2, Appl
35	65	9.5	797	2	US-08-362-240A-2	Sequence 2, Appl
36	65	9.5	797	5	PCT-US95-10245-2	Sequence 2, Appl
37	65	9.5	1012	1	US-07-944-943-2	Sequence 2, Appl
38	65	9.5	1012	1	US-08-219-262B-2	Sequence 2, Appl
39	65	9.5	1012	1	US-08-219-262B-14	Sequence 2, Appl
40	65	9.5	1012	3	US-09-031-655-2	Sequence 14, Appl
41	65	9.5	1012	3	US-09-031-655-14	Sequence 14, Appl
42	64.5	9.4	332	4	US-09-331-581-23	Sequence 23, Appl
43	64	9.3	380	1	US-08-416-478A-6	Sequence 6, Appl
44	64	9.3	380	2	US-08-474-988B-6	Sequence 6, Appl
45	64	9.3	380	2	US-08-394-442B-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1
US-08-330-272-4
Sequence 4, Application US/08330272
Patent No. 5985598

GENERAL INFORMATION:

APPLICANT: Russo et al
TITLE OF INVENTION: TCR-1 Gene and Related
TITLE OF INVENTION: Methods and Compositions
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/330,272
FILING DATE:
CLASSIFICATION: 530

ATTORNEY/AGENT INFORMATION:

NAME: MISTROCK, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6754-027
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 790-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 4:

SEQUENCE CHARACTERISTICS:

LENGTH: 108 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-330-272-4

Query Match 25.8%; Score 176.5; DB 2; Length 108;
Best Local Similarity 33.6%; Pred. No. 6.5e-14;

Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;

QY 1 MASEASVRLGVPGRMTWQR---VRFNPSRREWARAGSGSYEP 57
DB 1 MAGE---DVGAPPHLWVHGEGYRDEYGRITVAVAVEERSFLIRAVQIOVPLGDANP 57


```

; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6754-027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 790-8864/9741
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 113 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; PCT-US95-13663-2

```

```

Query Match      21.8%; Score 149; DB 5; Length 113;
Best Local Similarity 29.2%; Pred. No. 1.4e-10;
Matches 33; Conservative 21; Mismatches 45; Indels 14; Gaps 1;

QY 13 PGRMTORGITDEDEGRVTWTVVFNPSRREARASQGRVPSITYHMKQAVHTRE 72
    ||| : ||| : : : : : : : : : : : : : : : : : : : : : : :
DB 14 PDRLMMEFYVLDKQAHMLPTITIKRQ-----LRLVLRREDVYVLR 59
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 73 LSSGOMPSPQLPAVMQLYPGKRYRAADSFWETADHGQIDSMQOLVYTOPE 125
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB 60 PMPTQIGHSLLPTIMQLYPDGRYSSDSFWRLVYHIKIDGVEDMLLELPD 112
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 5
US-08-216-276A-33
; Sequence 33, Application US/08216276A
; Patent No. 5595912
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SNYDER, DAVID
; TITLE OF INVENTION: SPECIFIC DNA AND RNA SEQUENCES
; TITLE OF INVENTION: ASSOCIATED WITH US 1BDV VARIANTS, VECTOR CARRYING DNA
; TITLE OF INVENTION: SEQUENCES, HOST CARRYING CLONED VECTOR, DEDUCED AMINO ACID
; TITLE OF INVENTION: SEQUENCES, VACCINE AND METHOD OF VACCINATION
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,276A
; FILING DATE: 23-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/083,784
; FILING DATE: 28-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/519,202
; FILING DATE: 04-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/227,311
; FILING DATE: 02-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelber, Steven B.
; REGISTRATION NUMBER: 30,073

```

```

; REFERENCE/DOCKET NUMBER: 2747-054-27 CIP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 33:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 476 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-216-276A-33

```

```

Query Match      10.1%; Score 69; DB 1; Length 476;
Best Local Similarity 23.8%; Pred. No. 3.8;
Matches 30; Conservative 18; Mismatches 50; Indels 28; Gaps 5;

QY 8 RLGVPPGLMTORGITDEDEGRVTWTVVFNPSRREAR-----ASQGRYYP 57
    ||| : ||| : : : : : : : : : : : : : : : : : : : : : : :
DB 162 RLQ-----LKLGGGADVNTGPMATFIKRFPHNPRDMLRPLNLPYLPNNAGROY-- 214
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 58 SITVHMQAVHTREL--LSSGOMPSPQLPAVMQLYPGKRYRAADSFWETADHGQIDSM 115
    ||| : ||| : : : : : : : : : : : : : : : : : : : : : : :
DB 215 ---HLAMASERFETPELESAYRAMEAASVDL-----FQSALSYFMLENGVITDM 265
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 116 EQLYLT 121
    ||| : ||| : : : : : : : : : : : : : : : : : : : : : : :
DB 266 ANPAL 271
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

RESULT 6
US-08-216-276A-31
; Sequence 31, Application US/08216276A
; Patent No. 5595912
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SNYDER, DAVID
; TITLE OF INVENTION: SPECIFIC DNA AND RNA SEQUENCES
; TITLE OF INVENTION: ASSOCIATED WITH US 1BDV VARIANTS, VECTOR CARRYING DNA
; TITLE OF INVENTION: SEQUENCES, HOST CARRYING CLONED VECTOR, DEDUCED AMINO ACID
; TITLE OF INVENTION: SEQUENCES, VACCINE AND METHOD OF VACCINATION
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
; STREET: 1755 S. Jefferson Davis Highway, Suite 400
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/216,276A
; FILING DATE: 23-MAR-1994
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/083,784
; FILING DATE: 28-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/519,202
; FILING DATE: 04-MAY-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/227,311
; FILING DATE: 02-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelber, Steven B.
; REGISTRATION NUMBER: 30,073
; REFERENCE/DOCKET NUMBER: 2747-054-27 CIP

```

TELECOMMUNICATION INFORMATION
TELEPHONE: (703) 413-3000
TELEFAX: (703) 413-2220
TELEX: 248655 OPAF UR
INFORMATION FOR SEQ ID NO: 31:
SEQUENCE CHARACTERISTICS:
LENGTH: 540 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
DS-08-216-276A-31

Query Match	10.1%;	Score 69;	DB 1;	Length 540;
Best Local Similarity	23.8%;	Pred. No. 4.5;		
Matches 30;	Conservative 18;	Mismatches 50;	Indels 28;	Gaps 5

```

QY      8  RLCPGPGMTLQIORGXIGYEDEEERTVTVVYVFNPSREAR-----ASCSREP 57
          |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      403  RLG-----LKLAGGAFDVNTGPNNAIFIRPFHNRDMRPLPLNTPYLPAGROX-- 455
          |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      58  SITVHLKQMAVHREL--LSSQGMFQLPAVMQLPGKRYRADSSFWIEIDQIOM 115
          |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      456  ----HMAAASEFKEIPELESAYRAMEAASVPL-----FGSALSVFMLEENGIVTDM 506
          |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
QY      116  EQLVLT 121
          |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      507  ANFALS 512
          |||:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 7
US-08-216-276A-19
Sequence 19, Application US/08216276A
Patent No. 5595912
GENERAL INFORMATION:
APPLICANT: YAKHARIA, VIKRAM
APPLICANT: SNIDER, DAVID
TITLE OF INVENTION: SPECIFIC DNA AND RNA SEQUENCES
TITLE OF INVENTION: ASSOCIATED WITH US IBDV VARIANTS, VECTOR CARRYING DNA
TITLE OF INVENTION: SEQUENCES, HOST CARRYING CLONED VECTOR, DEDICED AMINO ACID
TITLE OF INVENTION: SEQUENCES, VACCINE AND METHOD OF VACCINATION
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESSES:
ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT,
ADDRESSEE: P. C.
STREET: 1755 S. Jefferson Davis Highway, Suite 400
CITY: Arlington
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/216,276A
FILING DATE: 23-MAR-1994
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/083,784
FILING DATE: 28-JUN-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/519,202
FILING DATE: 04-MAY-1990
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/227,311
FILING DATE: 02-AUG-1988
ATTORNEY/AGENT INFORMATION:
NAME: Kelbet, Steven B.
REGISTRATION NUMBER: 30,073
REFERENCE/DOCKET NUMBER: 2747-054-27 CIP
TELECOMMUNICATION INFORMATION:

```

: TELEPHONE: (703) 413-3000
:
: TELEFAX: (703) 413-2220
:
: TELEX: 248855 OPAT UR
:
: INFORMATION FOR SEQ. ID NO.: 19
:
: SEQUENCE CHARACTERISTICS:
:
:   LENGTH: 1012 amino acids
:   TYPE: amino acid
:   TOPOLOGY: linear
:
: MOLECULE TYPE: protein
:
US-08-216-276A-19

```

Query Match 10.1%; Score 69; DB 1; length 1012;
Best Local Similarity 23.8%; Pred. No. 11;
Matches 30; Conservative 18; Mismatches 50; Indels 28; Gaps 5;

```

QY      8 RLPFGRLIOISPGIIEDEEGRTWVVVVVFSPSSRMAR-----ASQSRREP 57
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      698 RLIG-----LTLAIPGAFIVNTGPNMAIFIKRFPNPRMDLPLYNLPYPNAGRY-- 750
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      58 SITTHIQMAVTHREL--LSSQMPFSQLPAYWQLYPGKRYRADSSFEWIDQIQISM 115
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      751 ----HLMASSEKREPELESIVRAMAASVDP-----PQSLSVFMWLEBNGIVDM 801
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY      116 EQLYIT 121
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      802 ANEALS 807
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 8
 US-08-219-262B-1
 Sequence 1, Application US/08219262B
 Patent No. 5788970
 GENERAL INFORMATION:
 APPLICANT: VAKHARIA, VIKRAM
 APPLICANT: SNYDER, DAVID B
 APPLICANT: MENDEL-WHERSAT, STEPHANIE A
 TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAI DISEASE VIRUS
 TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
 TITLE OF INVENTION: THEBON
 NUMBER OF SEQUENCES: 15
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: OHLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
 STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
 CITY: ARLINGTON
 STATE: VIRGINIA
 COUNTRY: USA
 ZIP: 22202
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC Compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/219,262B
 FILING DATE: 29-MAR-1994
 CLASSIFICATION: 435
 ATTORNEY/AGENT INFORMATION:
 NAME: OHLON, NORMAN F
 REGISTRATION NUMBER: 24,618
 REFERENCE/DOCKET NUMBER: 2747-047-27
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 413-3000
 TELEFAX: (703) 413-2220
 TELEX: 248855 OPAT UR
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1012 amino acids
 TYPE: amino acid
 STRANDEDNESS: unknown
 TOPOLOGY: unknown
 MOLECULE TYPE: protein
 ORIGINAL SOURCE:

ORGANISM: Infectious bursal disease virus
STRAIN: GLS
US-08-219-262B-1

Query Match 10.1%; Score 69; DB 1; Length 1012;
Best Local Similarity 23.8%; Pred. No. 11;

Matches 30; Conservative 18; Mismatches 50; Indels 28; Gaps 5;

QY 8 RLGVPPGRIMIORPGIYDEEGRTWTVVVRNPSRREMAR-----ASQSRYP 57
DB 698 RLG-----LKLGGPAGFDVNTGPNWATFIKRPHPNRDMDRLPYLNLPLYPPNAGROY-- 750
QY 58 STYVHLMQMAVHTREL--LSSGQPFSQLPAWOLYPERKYRAADSSFEIADHQIDSM 115
DB 751 -----HLMMAASEFKETPELESAYVAMEAASVDPL-----FQSALSVFWMLENGIYVDM 801
QY 116 EQLVLT 121
DB 802 ANRALS 807

RESULT 9
US-08-219-262B-9
Sequence 9, Application US/08219262B
Patent No. 5788970

GENERAL INFORMATION:

APPLICANT: VAKHARIA, VIKRAM

APPLICANT: SNYDER, DAVID B

TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS

TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT

STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

APPLICATION NUMBER: US/08/219,262B

FILING DATE: 29-MAR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 2747-047-27

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 1012 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: protein

ORIGINAL SOURCE:

ORGANISM: Infectious bursal disease virus

STRAIN: 002-73

US-08-219-262B-9

Matches 31; Conservative 16; Mismatches 51; Indels 28; Gaps 5;

QY 8 RLGVPPGRIMIORPGIYDEEGRTWTVVVRNPSRREMAR-----ASQSRYP 57
DB 698 RLG-----LKLGGPAGFDVNTGPNWATFIKRPHPNRDMDRLPYLNLPLYPPNAGROY-- 750
QY 58 STYVHLMQMAVHTREL--LSSGQPFSQLPAWOLYPERKYRAADSSFEIADHQIDSM 115
DB 751 -----HLMMAASEFKETPELESAYVAMEAASVDPL-----FQSALSVFWMLENGIYVDM 801
QY 116 EQLVLT 121
DB 802 ANRALS 807

RESULT 10
US-08-219-262B-12

Sequence 12, Application US/08219262B
Patent No. 5788970

GENERAL INFORMATION:

APPLICANT: VAKHARIA, VIKRAM

APPLICANT: SNYDER, DAVID B

TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS

TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MATER & NEUSTADT

STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR

CITY: ARLINGTON

STATE: VIRGINIA

COUNTRY: USA

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

APPLICATION NUMBER: US/08/219,262B

FILING DATE: 29-MAR-1994

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: OBLON, NORMAN F

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 2747-047-27

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703) 413-3000

TELEFAX: (703) 413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 1012 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-219-262B-12

Query Match 10.1%; Score 69; DB 1; Length 1012;
Best Local Similarity 23.8%; Pred. No. 11;

Matches 30; Conservative 18; Mismatches 50; Indels 28; Gaps 5;

QY 8 RLGVPPGRIMIORPGIYDEEGRTWTVVVRNPSRREMAR-----ASQSRYP 57
DB 698 RLG-----LKLGGPAGFDVNTGPNWATFIKRPHPNRDMDRLPYLNLPLYPPNAGROY-- 750
QY 58 STYVHLMQMAVHTREL--LSSGQPFSQLPAWOLYPERKYRAADSSFEIADHQIDSM 115
DB 751 -----HLMMAASEFKETPELESAYVAMEAASVDPL-----FQSALSVFWMLENGIYVDM 801
QY 116 EQLVLT 121

Db 802 ANFALS 807

```

RESULT 11
US-09-031-655-1
: Sequence 11, Application US/09031655
: Patent No. 6017759
:
: GENERAL INFORMATION:
: APPLICANT: VAKHARIA, VIKRAM
: APPLICANT: SANDER, DAVID B
: APPLICANT: MENDEL-WHERSAT, STEPHANIE A
: TITLE OF INVENTION: CHEMERIC INFECTIOUS BURSAL DISEASE VIRUS
: TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
: TITLE OF INVENTION: THEREON
: NUMBER OF SEQUENCES: 15
:
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT
: STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
: CITY: ARLINGTON
: STATE: VIRGINIA
: COUNTRY: USA
: ZIP: 22202
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/031,655
: FILING DATE:
:
: CLASSIFICATION:
: PRIORITY APPLICATION DATA:
: APPLICATION NUMBER: US 08/219,262
: FILING DATE: 29-MAR-1994
:
: ATTORNEY/AGENT INFORMATION:
: NAME: OBLON, NORMAN F
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 2747-047-27
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703) 413-3000
: TELEFAX: (703) 413-2220
: TELE: 248855 OPAT DR
:
: INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1012 amino acids
: TYPE: amino acid
: STRANDEDNESS: unknown
: TOPOLOGY: unknown
: MOLECULE TYPE: protein
: ORIGINAL SOURCE:
: ORGANISM: Infectious bursal disease virus
: STRAIN: GLS
:
: US-09-031-655-1

```

RESULT 12
US-09-031-655-9
; Sequence 9, Application US/09031655
; Patient No. 6017759
; GENERAL INFORMATION:
; APPLICANT: VAKHARIA, VIKRAM
; APPLICANT: SNYDER, DAVID B
; APPLICANT: MENDEL-WHERSAT, STEPHANIE A
; TITLE OF INVENTION: CHEMICAL INFECTIOUS BURSAL DISEASE VIRUS
; TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
; TITLE OF INVENTION: THERCON
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, WAIR & NEUSTADT
; STREET: 175 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: USA
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,655
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/219,262
; FILING DATE: 29-MAR-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2747-047-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-2200
; TELEFAX: (703) 413-2220
; TELEX: 248855 OPAIT UR
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1012 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Infectious bursal disease virus
; STRAIN: 002-73
; US-09-031-655-9

RESULT 13
US-09-031-655-12
; Sequence 12, Application US/09031655
; Patent No. 6017759

us-09-526-329-39.ra

```

1 GENERAL INFORMATION:
2 APPLICANT: VAKHARIA, VIKRAM
3 APPLICANT: SNYDER, DAVID B
4 APPLICANT: MENDEL-WHIGRETT, STEPHANIE A
5 TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
6 TITLE OF INVENTION: CHIMERIC INFECTIOUS BURSAL DISEASE VIRUS
7 TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
8 NUMBER OF SEQUENCES: 15
9
10 CORRESPONDENCE ADDRESS:
11 ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MATER & NEUSTADT
12 STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
13 CITY: ARLINGTON
14 STATE: VIRGINIA
15 COUNTRY: USA
16 ZIP: 22202
17
18 COMPUTER READABLE FORM:
19 MEDIUM TYPE: Floppy disk
20 COMPUTER: IBM PC compatible
21 OPERATING SYSTEM: PC-DOS/MS-DOS
22 SOFTWARE: PatentIn Release #1.0, Version #1.30
23 CURRENT APPLICATION DATA:
24 APPLICATION NUMBER: US/09/031,655
25 FILING DATE:
26
27 CLASSIFICATION:
28 PRIOR APPLICATION DATA:
29 APPLICATION NUMBER: US 08/219,262
30 FILING DATE: 29-MAR-1994
31 ATTORNEY/AGENT INFORMATION:
32 NAME: OBLON, NORMAN F
33 REGISTRATION NUMBER: 24,618
34 REFERENCE/DOCKET NUMBER: 2747-047-27
35 TELECOMMUNICATION INFORMATION:
36 TELEPHONE: (703) 413-3000
37 TELEFAX: (703) 413-2220
38 TELE: 248855 OPAT UR
39 INFORMATION FOR SEQ. ID NO. 12:
40 SEQUENCE CHARACTERISTICS:
41 LENGTH: 1012 amino acids
42 TYPE: amino acid
43 TOPOLOGY: linear
44
45 MOLECULE TYPE: protein
46
47 US-09-031-655-12

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040

```

```

Query Match 9.9%; Score 68; Db 2; Length 1013;
Best Local Similarity 23.1%; Pred. No. 14;
Matches 27; Conservative 19; Mismatches 61; Indels 10; Gaps 4.

QY      8  RLGYPPERLMIQPGITDEBGRWTWYVVFNFPSRRREMARASQ-GSRYPESTVTHLQM 66
      111:  : 111:  : 111:  : 111:  : 111:  : 111:  : 111:  : 111:  :
Db      699  RLKMG-----LPGAVADINTGNMATEFKKRPHPNPDMDRLPLNLPPLPAGRGHLL 753
      111:  : 111:  : 111:  : 111:  : 111:  : 111:  : 111:  : 111:  :

QY      67  AVHRELLSGOM--PFSOLPAWMLYGRKRYRAADSFWELIADHGQIDSMEDVLT 121
      111:  : 111:  : 111:  : 111:  : 111:  : 111:  : 111:  : 111:  :
Db      754  ALASSEKETPELEDVARADAANADP--LFRSALQVEMWLEENGIVTDMANFALS 808
      111:  : 111:  : 111:  : 111:  : 111:  : 111:  : 111:  : 111:  :

RESULT 15
US-08-219-262b-7
; Sequence 7, Application US/08219262B
; Patent No. 5788970
;
GENERAL INFORMATION:
;
APPLICANT: YAKHARIA, VIRAM
APPLICANT: SNYDER, DAVID B
APPLICANT: MENDEL-WHERSAT, STEPHANIE A
TITLE OF INVENTION: CHIMERIC INECTIONS BURSAL DISEASE VIRUS
TITLE OF INVENTION: CDNA CLONES, EXPRESSION PRODUCTS AND VACCINES BASED
TITLE OF INVENTION: THEREON
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESS: OBLON, SPIYAK, MCCLELLAND, MAIER & NEUSTADT
STREET: 1755 S. JEFFERSON DAVID HIGHWAY, FOURTH FLOOR
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: USA
ZIP: 22202
;
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/219,262B
FILING DATE: 29-MAR-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

```

```

; NAME: OBLON, NORMAN F
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 2747-047-27
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 413-3000
; TELEFAX: (703) 413-2220
; TELEX: 24885 OPAT UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1012 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: unknown
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Infectious bursal disease virus
; STRAIN: 52/70
; US-08-219-262B-7

```

Query Match 9.9%; Score 67.5; DB 1; Length 1012;

Best local similarity 23.9%; Pred. No. 16; Mismatches 28; Indels 41; Gaps 5;

```

Matches 26; Conservative 14; Mismatches 28; Indels 41; Gaps 5;

QY 8 RLGVPGRIMIQRPQIYDEDEGRITVYVVFNPSSREMARASOGSKYEPSTIVHLMOMA 67
   || | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 698 RLG-----LKLGGAGADVNTGPNMNFIRKRPHPMDR----- 733

QY 68 VHTRELLSSGOMPSSLPVWQLYP--GRKYRAADSSFEWELADHGQIDS 114
   : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 734 -----LpyLNLp---YLPPNAGRQYHLMAAS--EFKDTPELES 767

```

Search completed: November 29, 2001, 04:03:37
Job time: 333 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 03:59:59 ; Search time 18.53 Seconds
(without alignments)
526.192 Million cell updates/sec

Title: US-09-526-329-39

Perfect score: 685

Sequence: 1 MASBASVRLGVPPGRLMTOR.....HGQIDSMQVLVYQPERKD 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	176.5	25.8	107	2	S78532
2	149	21.8	114	2	I38286
3	79.5	11.6	447	2	S18339
4	79.5	11.6	603	2	S03020
5	79.5	11.6	603	2	S03020
6	79	11.5	238	2	C83240
7	75.5	11.0	476	2	E83796
8	74.5	10.9	436	2	S06884
9	73.5	10.7	261	2	J00137
10	73	10.7	463	2	G83175
11	72.5	10.6	424	2	B25063
12	72.5	10.6	513	2	T34689
13	72.5	10.6	1414	2	B70674
14	70	10.2	367	2	C82391
15	70	10.2	830	2	T07824
16	70	10.2	861	1	S34730
17	69.5	10.1	357	2	J01647
18	69.5	10.1	363	2	T44150
19	69.5	10.1	433	2	T43961
20	69.5	10.1	1927	2	A59236
21	69	10.1	363	2	T36024
22	69	10.1	1012	1	GNXSAU
23	69	10.1	1747	2	T43162
24	69	9.9	1013	2	S32214
25	67.5	9.9	762	2	S60415
26	67.5	9.9	1012	1	GNX552
27	67.5	9.9	1148	2	H86674
28	67	9.8	1484	2	I40847
29	66.5	9.7	536	2	S30973

Query Match Length DB ID Description
MTCP-1 protein sp1
T cell leukemia/ly
vird2 protein - Ag
vird2 protein - Ag
dopamine beta-mono
probable transcrip
two-component sens
vird2 protein - Ag
hypothetical 30.1K
hypothetical metallo-o
probable iron-sulf
probable mltb prot
N-ethylmaleimide r
1.4-alpha-glucan b
SHL1 protein - hum
hypothetical prote
embryonic muscle m
conserved hypothet
genome polyprotein
vitellogenin - gyp
hypothetical prote
dolichyl-phosphate
genome polyprotein
hypothetical prote
succinate dehydrog
minor tail protein

30	66.5	9.7	1209	2	H85839	probable regulator
31	66	9.6	237	2	C83183	probable transcrip
32	66	9.6	363	2	T37072	hypothetical prote
33	66	9.6	536	2	D83622	arylsulfatase PA01
34	66	9.6	993	1	GNX598	genome polyprotein
35	66	9.6	1001	2	C86181	hypothetical prote
36	66	9.6	1012	1	GNX519	genome polyprotein
37	66	9.6	1012	1	GNX519	genome polyprotein
38	66	9.6	1170	2	H71295	probable DNA polym
39	66	9.6	1175	2	T25634	hypothetical prote
40	65.5	9.6	327	2	T15594	hypothetical prote
41	65.5	9.6	350	2	D83102	xenobiotic reducta
42	65.5	9.6	764	2	D84847	probable receptor-
43	65	9.5	286	2	S48201	licheninase (EC 3.
44	65	9.5	352	2	S76078	hypothetical prote
45	65	9.5	406	2	C83867	Xaa-Pro dipeptidase

ALIGNMENTS

RESULT 1
S78532
MTCP-1 protein splice form B1 - human
C:Species: Homo sapiens (man)
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000
C:Accession: S78532
R:Stern, M.H.; Soulier, J.; Rosenzweig, M.; Nakahara, K.; Canki-Klajn, N.; Aurias, A.
Oncogene 8, 2475-2483, 1993
A>Title: MTCP-1: a novel gene on the human chromosome Xq28 translocated to the T cell
A:Reference number: 138045; MUID:93368950
A:Accession: S78532
A:Molecule type: DNA
A:Residues: 1-107 <STE>
A:Cross-references: EMBL:Z24459; NID:g2252491; PIDN:CAA80828.1; PID:g2252492
C:Genetics:
A:Gene: MTCP-1
A:Introns: 35/3; 92/3
C:Keywords: alternative splicing; T-cell proliferation

Query Match 25.8; Score 176.5; DB 2; Length 107;
Best Local Similarity 33.6%; Pred. No. 1.9e-11;
Matches 43; Conservative 17; Mismatches 43; Indels 25; Gaps 4;
QY 1 MASBASVRLGVPPGRLMTORPGIYDEDEGRMTVTVV---VRENPSPREARASGSRVPE 57
DB 1 MAGE--DVGAPDHLYWROGIRYRTWVAVVEETSLRLARVQDIQVPLGDAAR 57
QY 58 SITVHLQMAVTRRELSSGOMPSPQPAVMQVYPRKRYRADSSFWETADHQIDSMQ 117
DB 58 S-----HLT-----SOLPLMQVLYPEERYMDNNSRLMOIHHLVAVGOE 98
QY 118 LVLTQPE 125
DB 99 LLLKLPPD 106
RESULT 2
I38286
T cell leukemia/lymphoma protein TCL1 - human
N:Alternate names: T-cell leukemia-related protein TCL1
C:Species: Homo sapiens (man)
C>Date: 09-Mar-1996 #sequence_revision 09-Mar-1996 #text_change 05-Nov-1999
R:Virgilio, L.; Narducci, M.G.; Isobe, M.; Billings, L.G.; Cooper, M.D.; Croce, C.M.;
Proc. Natl. Acad. Sci. U.S.A. 91, 12530-12534, 1994
A>Title: Identification of the TCL1 gene involved in T-cell malignancies.
A:Reference number: I38286; MUID:95107991
A:Accession: I38286
A:Molecule type: mRNA
A:Residues: 1-114 <RES>
A:Cross-references: EMBL:X82240; NID:g624960; PIDN:CAA57708.1; PID:g624961

Db 377 TDCKTQALPPSGIHIFASDLHFLGKRVYTVLR--DGREWEIYNODNHYSF----- 428

QY 63 LQMAVHTRELSSGQMPFSQLPAVMOLYPCR-----KYRADSPFWEIADHGOIDSE 116

Db 429 -----HFOFI-----RMLKRVSVHVGDDVLTSTCTVTEER---ELAIVGCGFGLTE 471

QY 117 QLVVTV 122

Db 472 EMCVNY 477

RESULT 6

C83240

probable transcription regulator PA3249 [Imported] - Pseudomonas aeruginosa (strain PA01

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000

C:Accession: C83240

R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br

adman, S.; Yuan, T.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Linn,

..; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho

A:Reference number: A82950; MUID:20437337

A:Accession: C83240

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-238 <STO>

A:Cross-references: GB:AE004747; GB:AE004091; NID:99949362; PIDN:AA06637.1; GSPDB:GN001

A:Experimental source: strain PA01

A:Genetics:

A:Gene: PA3249

Query Match 11.5%; Score 79; DB 2; Length 238;

Best Local Similarity 23.9%; Pred. No. 0.75;

Matches 28; Conservative 17; Mismatches 46; Indels 26; Gaps 3;

QY 5 ASVRLGVPPRLMIGRPGLYEDEGRFTWTVVVRNPSRRMARASQGSRYEPSITVHL 64

Db 42 ATTRITRLREALIOLESQGLTYREE-----RGMFVSPERLAVNPLVRSHF 87

QY 65 QMA-----VHTRELSSGQMPF-----QLPAVMOLYPCRKYRADSPFWEIADH 109

Db 88 AMVSEGRVPADEVLSARQMPAAVCELELPALSSVYIRARVRDGLVTVYEH 144

RESULT 7

E83796

two-component sensor histidine kinase BH1173 [Imported] - Bacillus halodurans (strain C-

C:Species: Bacillus halodurans

C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 31-Dec-2000

C:Accession: E83796

R:Takeuchi, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira

Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and

A:Reference number: A83650; MUID:20263314

A:Accession: E83796

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-476 <STO>

A:Cross-references: GB:AP001511; GB:BA000004; NID:910173727; PIDN:BA004892.1; GSPDB:GN00

A:Experimental source: strain C-125

A:Genetics:

A:Gene: BH1173

Query Match 11.0%; Score 75.5; DB 2; Length 476;

Best Local Similarity 27.7%; Pred. No. 4.1;

Matches 36; Conservative 16; Mismatches 53; Indels 25; Gaps 7;

QY 19 ORRGTYE-DEGRFTWTVVVRNPSRR-MAASQGSRYEPSITVHLQMAVHTRELSS 76

Db 115 QSRGHTSLNGATITTYVTRKNSGREGVLTISYMDTYRDSMVNRILMKLYTL-LTSS 173

QY 77 GQMPFSQLPAVMQ-----LYPGKTRAADSF-----WEIAD-----HGOIDSE 117

Db 174 A---LSLTPALMKHYLRQPLILLGNRLQIADNRWKEPFKMGDEDFOKLSNOFERRQ 230

QY 118 LVLTQYPERK 127

Db 231 NLVRYDOSQK 240

RESULT 8

S06884

virD protein - Agrobacterium rhizogenes plasmid pRiA4b

C:Species: Agrobacterium rhizogenes

C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 08-Oct-1999

C:Accession: S06884

R:Hirayama, T.; Muranaka, T.; Ohkawa, H.; Oka, A.

Mol. Gen. Genet. 213, 229-237, 1988

A:Title: Organization and characterization of the virD genes from Agrobacterium rhiz

A:Reference number: S06881; MUID:89039712

A:Accession: S06884

A:Molecule type: DNA

A:Residues: 1-436 <HIR>

A:Cross-references: EMBL:X12867; NID:938995; PIDN:CAA31351.1; PID:939000

A:Genetics:

A:Gene: virD2

A:Genome: plasmid

Query Match 10.9%; Score 74.5; DB 2; Length 436;

Best Local Similarity 25.6%; Pred. No. 4.7;

Matches 34; Conservative 17; Mismatches 43; Indels 39; Gaps 8;

QY 11 VPP-----GRLMIGRPGLYE-----DEGRFTWTV--VVYRF-----NPSRRMA-- 47

Db 49 LPDQIHELARSWQETGYDESQDERQDELTHITIVSPAGTSQAAMASRRMAE 108

QY 48 ---RASQGSRYEPSITVHL-----WQMAVHTRELSSGQMPFSQLPAVMOLYPCRKYRA 98

Db 109 MEGSAGGSGSTNYLTAFHIDRPHLHVAVNRRLGHGWLKIS-----RRHPQLNYDA 162

QY 99 AQSFWETA-DHG 110

Db 163 LRINMAETSLRHG 175

RESULT 9

J00137

hypothetical 30.1k protein - Pseudomonas aeruginosa

C:Species: Pseudomonas aeruginosa

C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 07-Jun-1996

C:Accession: J00137

R:Kato, J.; Chu, L.; Kitano, K.; Devault, J.D.; Kimbara, K.; Chakrabarty, A.M.; Misra

Gene 84, 31-38, 1989

A:Title: Nucleotide sequence of a regulatory region controlling alginate synthesis in

A:Reference number: J00132; MUID:90108714

A:Accession: J00137

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-261 <KAT>

A:Note: 3-Met could also be the initiator

A:Genetics:

A:Start codon: GTG

Query Match 10.7%; Score 73.5; DB 2; Length 261;

Best Local Similarity 27.0%; Pred. No. 3.2;

Matches 30; Conservative 11; Mismatches 41; Indels 29; Gaps 6;

QY 13 PERLMIORPGIYE-----DEGRFTWTVVVRNPSRRMARAS-----QGSREP 57

Db 51 PKQMRRLRPGRMESPORSGQERWRAMLRQVSRILRVSPQAMPPEVSPQSPQAMLQASR--P 108

OY 58 SITVHLQMVAHTRELLSSGQMPFSQLPAWV-----QLYGKRYRADSSP 103
 Db 109 RVSPHAMPAA-----WLRASRLRFS--PRAMPVSPQASBPMLARSLNRP 152

RESULT 10
 G83175
 probable metallo-oxidoreductase PA3768 [Imported] - Pseudomonas aeruginosa (strain PA01)

C:Species: Pseudomonas aeruginosa
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 31-Dec-2000
 C:Accession: G83175
 R:Stover, C.K.; Pham, X.Q.; Ewain, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Berman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.K.; Kas, A.; Laidig, K.; Lim, J.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
 A:Reference number: A82950; MUID:20437337
 A:Accession: G83175
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-463 <STO>
 A:Cross-references: GB:AE004795; GB:AE004091; NID:g9949931; PIDN:AA07155.1; GSPDB:GN001
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA3768

Query Match 10.7%; Score 73; DB 2; Length 463;
 Best local similarity 24.3%; Pred. No. 7.2;
 Matches 35; Conservative 18; Mismatches 55; Indels 36; Gaps 8;

OY 1 MASEASVRLGVPGRMTIQRRGI-----YEDEGRVTWTV-----YRF 39
 Db 12 LAGIAVVGAGAGATMLARQVAGETDELTAAPLDELTPGESSPALANGGCGEVEL 71
 OY 40 NPSRRMARAQSGRYSPTVHLQMVAHTRELLSSGQMPFSQLPAWOLYPRGK 95
 Db 72 RANQGMELVRFNTNLDPEPTIH-WH--GIRLPEDMGVPYISQPPV---QPGSFYQ 124
 OY 96 YRAADS-SFWEIADHGQIDSMEO 118
 Db 125 FKTDAGSYWY---HPLMSSEQL 145

RESULT 11

B25063

hypobacterial virD2 protein - Agrobacterium tumefaciens plasmids

C:Species: Agrobacterium tumefaciens

C:Date: 31-Mar-1988 #sequence_revision 31-Mar-1988 #text_change 08-Oct-1999

C:Accession: B25063

R:Jayaswal, R.K.; Veluthambi, K.; Gelvin, S.B.; Slightom, J.L.

J. Bacteriol. 169, 5035-5045, 1987

A:Title: Double-stranded cleavage of T-DNA and generation of single-stranded T-DNA molecules

A:Reference number: A91846; MUID:88032822

A:Accession: B25063

A:Molecule type: DNA

A:Residues: 1-424 <DNA>

A:Cross-references: EMBL:M17989; NID:g142288; PIDN:AAA22114.1; PID:g142290

A>Note: plasmid pTiA6K

R:Yanovsky, M.F.; Porter, S.G.; Young, C.; Albright, L.M.; Gordon, M.P.; Nester, E.W.

Cell 47, 471-477, 1986

A:Title: The virD operon of Agrobacterium tumefaciens encodes a site-specific endonuclease

A:Reference number: A90886; MUID:87028239

A:Accession: B25063

A:Molecule type: DNA

A:Residues: 1-424 <DNA>

A>Note: plasmid pTiA6K

C:Genetics:

A:Gene: virD2

A:Genome: plasmid

C:Keywords: endonuclease

Query Match 10.6%; Score 72.5; DB 2; Length 424;
 Best local similarity 24.1%; Pred. No. 7.3;
 Matches 38; Conservative 24; Mismatches 49; Indels 47; Gaps 10;

OY 11 VPPGRRL-----MIQRPIY-----DEGRWTWTV-----NPSRRMARA 49
 Db 49 VPPDIRELAQSVTENGITDESDDDRODILTHIVSPAGTDQTAAYEASREMAE 108
 OY 50 SGG-----RYEBSITVHL-----WQMAVHTRELLSSGQMPFSQLPAWOL-YGKRYR 97
 Db 109 MFGSGGGGRVNLTAHYDRDHPHLHVYVNRRELLGHGWLKISRHP--QLNYGDLKK 166
 OY 98 AADSEFWEIADHGQI-----DSMEDLVLTQPERK 127
 Db 167 MAEISL-----RHGIVDAISRERKRIAPERITTAHERR 200

RESULT 12

T34689

probable iron-sulfur binding oxidoreductase - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999

C:Accession: T34689

R:Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, May 1998

A:Reference number: Z21553

A:Accession: T34689

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-513 <HAR>

A:Cross-references: EMBL:AL023517; PIDN:CAA18985.1; GSPDB:GN00070; SCOEDB:SC1B5.11C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC1B5.11C

Query Match 10.6%; Score 72.5; DB 2; Length 513;
 Best local similarity 32.8%; Pred. No. 9.2;
 Matches 19; Conservative 7; Mismatches 19; Indels 13; Gaps 2;

OY 11 VPPGRRLTORPG-----IYDEBGR-----TWVVVVRFPNPSRRMARAQSGSRV 55
 Db 436 LPPGEGAVVVRAGGRLAVYDEBGAALHAVSPRCHGCLVDFNAERAMECPGSGSRF 493

RESULT 13

B70674

probable mbtB protein - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 03-Nov-2000

C:Accession: B70674

R:Coile, S.T.; Broesch, R.; Parkhill, J.; Garner, T.; Churcher, C.; Harris, D.; Gordon

Rajandream, M.A.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Najandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome

A:Reference number: A70500; MUID:98295987

A:Accession: B70674

A:Molecule type: DNA

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Residues: 1-1414 <COL>

A:Cross-references: GB:B81371; GB:AL123456; NID:g3261669; PIDN:CAB03756.1; PID:g16573

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: mbtB

C:Superfamily: Mycobacterium tuberculosis mbtB protein; acetate--CoA ligase homology;

C:Keywords: carrier protein

E:5-74/Domain: acyl carrier protein homology <ACP>

E:598-1037/Domain: acetyl-CoA ligase homology <ACL>

E:1060-1131/Domain: acyl carrier protein homology <ACPI>

E:1189-1400/Domain: oleoyl-[acyl-carrier protein] hydrolase homology <ACPH>

Query Match 10.6%; Score 72.5; DB 2; Length 1414;

Best Local Similarity 26.1%; Pred. No. 32;

Matches 30; Conservative 9; Mismatches 55; Indels 21; Gaps 2;

QY 5 ASVRLGVP-----PGRLLIQPGIYEDGRTWVTVVVFNPSPR 44

DB 862 ASVPGVGFPPNACRVVADSGDDCPDWVAGELWVSGRGANGYRGRELTAERFVEHGR 921

QY 45 EVARASGSRREPSITVHLMQMAVHTRELLSSGOMFSQLPRAWQLYPGKTRAA 99

DB 922 TWYRTGLARIWHDGTLFVGRADH-RVKISGYRELGEIRALQRLPGVHAAA 975

RESULT 14

C82391

N-ethylmaleimide reductase VCA0993 [Imported] - Vibrio cholerae (strain N16961 serogroup

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 02-Feb-2001

C:Accession: C82391

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;

Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, H.

1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: AB2035; MUID:20406833

A:Accession: C82391

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-367 <HEI>

A:Cross-references: GB:AE004426; GB:AE003853; NID:99658431; PIDN:AAF96889.1; GSPDB:GN004

C:Genetics:

A:Gene: VCA0993

A:Map position: 2

C:Superfamily: NADPH dehydrogenase chain CYE2

Query Match 10.2%; Score 70; DB 2; Length 367;

Best Local Similarity 24.8%; Pred. No. 11;

Matches 26; Conservative 18; Mismatches 33; Indels 28; Gaps 7;

QY 21 PGIEDEGRWTVTVVFNPSRREWARASGSRREPSITVHLMQMAVHTRELLSSGOMP 80

DB 71 PGVITDEQTSQKTYT-----QAVKQGA---AMFCQLHWGRVSHPFQKQLP 117

QY 81 FSQLPAAWQLYPGKRYRAADSSFEIAD--HGO-IDSMQVLYT 121

DB 118 IA--PSA-----LKPVEYQW-IADQGNQGMVDCVEPRAMT 151

RESULT 15

T07824

1,4-alpha-glucan branching enzyme (EC 2.4.1.18) I (clone sbe17) - potato (fragment)

C:Species: Solanum tuberosum (potato)

C:Date: 14-May-1999 #sequence_revision 14-May-1999 #text_change 21-Jul-2000

C:Accession: T07824

R:Khoshnoodi, J.; Blennow, A.; Ek, B.; Raak, L.; Larsson, H.

Eur. J. Biochem. 242, 148-155, 1996

A:Title: The multiple forms of starch branching enzyme I in Solanum tuberosum.

A:Reference number: Z16155; MUID:97112484

A:Accession: T07824

A:Status: preliminary; translated from GB/EMBL/DBD

C:Keywords: glycogen/starch biosynthesis; glycosyltransferase; hexosyltransferase

Query Match 10.2%; Score 70; DB 2; Length 830;

Best Local Similarity 24.2%; Pred. No. 30;

Matches 31; Conservative 17; Mismatches 34; Indels 46; Gaps 8;

QY 17 WQRPGLYEDGRTWTVVFNPSRREWARA-SQSRREPSITVHLMQMAVHTRE--- 72

DB 545 WIDFP-----REGNNW-----SYDKCRQWNLADSEHLRYK---FNNAFDRAMSLDEKE 591

QY 73 -LSSGOMPFS-----OLPAAWQLYPGR-----KTRA-DSSF 103

DB 592 SFLASGKQIVSSMDDDNKVVFERNGLVYFVFNHKNYEGYKVCDFGKTRVALDSDA 651

QY 104 WEIADHQ 111

DB 652 WEFGRGR 659

Search completed: November 29, 2001, 04:04:10

Job time: 251 sec

Fri Nov 30 11:03:11 2001

us-09-526-329-39.rpr

Page 6

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 29, 2001, 04:03:14 ; Search time 9.89 Seconds

(without alignments)
474,530 Million cell updates/sec

Title: US-09-526-329-39

Perfect score: 685
Sequence: 1 MASEASVRLGVPGRLWIOR.....HGQIDSMELVLTQPEKRD 128

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	685	100.0	128	1	TCLB_HUMAN
2	178.5	26.1	116	1	TBL1_MOUSE
3	178.5	26.1	120	1	TBL4_MOUSE
4	178	26.0	107	1	MTC2_MOUSE
5	176.5	25.8	107	1	MTC2_HUMAN
6	175.5	25.6	122	1	TBL3_MOUSE
7	169	24.7	116	1	TCL4_MOUSE
8	163.5	23.9	121	1	TBL5_MOUSE
9	158.5	23.1	117	1	TBL2_MOUSE
10	149	21.8	114	1	TCLA_HUMAN
11	79.5	11.6	447	1	VID2_AGR5
12	79.5	11.6	603	1	DOPO_HUMAN
13	74.5	10.9	436	1	VID2_AGRH
14	72.5	10.6	424	1	VID2_AGR6
15	70	10.2	861	1	GLGH_SOUTU
16	69.5	10.1	366	1	VUL_HSV6U
17	69.5	10.1	775	1	LYS4_EHENT
18	69.5	10.1	1411	1	Y297_HUMAN
19	69	10.1	1012	1	POLS_IBDVA
20	67.5	9.9	762	1	PMT4_YEAST
21	67.5	9.9	1012	1	POLS_IBDVS
22	67	9.8	125	1	DHSC_COXBU
23	67	9.8	3067	1	CAIC_MOUSE
24	66.5	9.7	595	1	VG28_BPMU5
25	66.5	9.7	1490	1	CRK7_HUMAN
26	66	9.6	326	1	TR44_BACR
27	66	9.6	535	1	ARS_PSBAB
28	66	9.6	993	1	POLS_IBDVP
29	66	9.6	1012	1	POLS_IBDVC
30	66	9.6	1012	1	POLS_IBDVS
31	65.5	9.6	1170	1	DP3A_TREPA
32	65.5	9.6	808	1	SYFA_SYPN7
33	65.5	9.6	877	1	SYA_THIFE

34	65	9.5	286	1	GUB_RHOMR	P45798 rhodothermu
35	65	9.5	1233	1	HCA_YA_OCTDO	P12659 octopus dof
36	64.5	9.4	295	1	TF_RAT	P42533 rattus norv
37	64.5	9.4	439	1	ACA8_MOUSE	Q90034 mus musculu
38	64.5	9.4	1148	1	MFD_ECOLI	P310958 escherichia
39	64	9.3	310	1	PARB_XYLEA	Q90053 xylella fas
40	64	9.3	529	1	LAC3_HUMAN	P18627 homo sapien
41	64	9.3	2410	1	MOK1_SCHPO	Q90097 schizosacch
42	63	9.2	380	1	KR15_HSV1	Q00097 ictaluriid h
43	63	9.2	684	1	CDK9_CAEEL	P46551 caenorhadi
44	63	9.2	1442	1	CPSA_HUMAN	Q10570 homo sapien
45	63	9.2	1444	1	CPSA_BOVIN	Q10569 bos taurus

ALIGNMENTS

RESULT 1
ID TCLB_HUMAN STANDARD; PRT; 128 AA.
AC 095968;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE T-CELL LEUKEMIA/LYMPHOMA PROTEIN 1B (TCLB ONCOGENE)
DE (SYNCTIOTROPHOBLAST-SPECIFIC PROTEIN) (SYN-1).
GN TCLB.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
CX NCBI_TaxID:9606;
RN 1;
RP SEQUENCE FROM N.A.
RA MEDLINE-99178995; PubMed1007617;
RX Bekarsky Y., Hallas C., Isobe M., Russo G., Croce C.M.;
RT "Abnormalities at 14q32.1 in T cell malignancies involve two
RT oncogenes";
RL Proc. Natl. Acad. Sci. U.S.A. 96:2949-2951(1999).
RN [2]
RP SEQUENCE FROM N.A.
RA Jiang B., Mendelson C.R.;
RT "A syncytiotrophoblast-specific gene Syn-1 cloned from human
RT syncytiotrophoblast subcloned cDNA library";
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- TISSUE SPECIFICITY: EXPRESSED IN A VARIETY OF TISSUES.
CC -!- DISEASE: ACTIVATED IN CHRONIC T-CELL LEUKEMIAS (T-CLL) CARRYING
CC REARRANGEMENTS OF THE 14Q32.1 REGION.
CC -!- SIMILARITY: BELONGS TO THE TCLB FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AF110466; AAD16997.1;
CC DR EMBL: AF110465; AAD16996.1;
CC DR EMBL: AF137027; AAD30130.1;
CC DR HSSB: P56278; IALX.
CC MIM: 603769;
CC DR InterPro: IPR002709; TCLB_MTCPL.
CC DR ProDom: PD015575; TCLB_MTCPL1.
CC KW Proto-oncogene; Chromosomal translocation.
CC
CC SEQUENCE 128 AA; 14846 MW; 68A2C3F40F78B4D CRC64;

Query Match 100.0%; Score 685; DB 1; Length 128;
Best Local Similarity 100.0%; Pred. No. 1e-65;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MASEASVRLGVPGRLWIORPGIYEDGRTWVTVVVRNPNRREMARAGSGSRKPEST 60

```

Db 1 MASEASVRLVPGRLMIOREIYEDDEGRTWTVVFNFSREMARASQSGREPSIT 60
    |||
Qy 61 VHLQMAVHTRELLSSGQMPFSQLPAYWQLYPGKRYRAADSFEIADHQIDSMEQVL 120
    |||
Db 61 VHLQMAVHTRELLSSGQMPFSQLPAYWQLYPGKRYRAADSFEIADHQIDSMEQVL 120
    |||
Qy 121 TYOPEPRD 128
    |||
Db 121 TYOPEPRD 128

```

RESULT 2

```

ID TLB1_MOUSE STANDARD: PRT; 116 AA.
AC P56840;

```

```

DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TLB1 PROTEIN.
GN TLB1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=20056259; PubMed=10588720;
RA Hallas C., Pekarsky Y., Itoyama T., Varnum J., Bichi R.,
RA Rothstein J.L., Croce C.M.;
RT "Genomic analysis of human and mouse Tc11 loci reveals a complex of
RT tightly clustered genes."
RL Proc. Natl. Acad. Sci. U.S.A. 96:14418-14423(1999).

```

```

CC -1- SIMILARITY: BELONGS TO THE Tc11 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
CC EMBL; AF195488; AAF12801.1; -.
CC MGD; MGI:1351604; Tc11b4.
CC InterPro; IPR002709; Tc11_MTCP1.
CC Prodom; PD015575; Tc11_MTCP1; 1.
CC Multigene family.
KW SEQUENCE 116 AA; 13432 MW; 5D2E08EBC7BE2A64 CRC64;

```

```

Query Match 26.1%; Score 178.5; DB 1; Length 116;
Best Local Similarity 34.2%; Pred. No. 3.5e-12;
Matches 41; Conservative 19; Mismatches 51; Indels 9; Gaps 2;

```

```

Qy 1 MASEASVRLVPGRLMIOREIYEDDEGRTWTVVFNFSREMARASQSGREPSIT 60
    |||
Db 1 MAAAADPGLPPLPYLVSVRLGIYEDDEHHRWIVANV-----ETSHSSHGNNRRTHVT 53
    |||
Qy 61 VHLQMAVHTRELLSSGQMPFSQLPAYWQLYPGKRYRAADSFEIADHQIDSMEQVL 120
    |||
Db 54 VHLQMAVHTRELLSSGQMPFSQLPAYWQLYPGKRYRAADSFEIADHQIDSMEQVL 120
    |||

```

```

RESULT 3
TLB1_MOUSE STANDARD: PRT; 120 AA.
AC P56844;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TLB1 PROTEIN.
GN TLB1b4.

```

```

OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RX MEDLINE=20056259; PubMed=10588720;
RA Hallas C., Pekarsky Y., Itoyama T., Varnum J., Bichi R.,
RA Rothstein J.L., Croce C.M.;
RT "Genomic analysis of human and mouse Tc11 loci reveals a complex of
RT tightly clustered genes."
RL Proc. Natl. Acad. Sci. U.S.A. 96:14418-14423(1999).

```

```

CC -1- SIMILARITY: BELONGS TO THE Tc11 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC -----
CC EMBL; AF195492; AAF12805.1; -.
CC MGD; MGI:1351604; Tc11b4.
CC InterPro; IPR002709; Tc11_MTCP1.
CC Prodom; PD015575; Tc11_MTCP1; 1.
CC Multigene family.
KW SEQUENCE 120 AA; 14123 MW; 7F58E0AEBFF339 CRC64;

```

```

Query Match 26.1%; Score 178.5; DB 1; Length 120;
Best Local Similarity 36.4%; Pred. No. 3.7e-12;
Matches 43; Conservative 13; Mismatches 53; Indels 9; Gaps 2;

```

```

Qy 6 SVR--LGVPRGLMIOREIYEDDEGRTWTVVFNFSREMARASQSGREPSITVH 62
    |||
Db 4 SVRPFMPFPCVCTRDYEDDEHGRGVAAKVT-----SHSPYCKLEKCVYVH 57
    |||
Qy 63 LMQMAVHTRELLSSGQMPFSQLPAYWQLYPGKRYRAADSFEIADHQIDSMEQVL 120
    |||
Db 58 LMQMTTLFQPSDPSLKTENFLPRTWRLESNNRYRGADAMHVRVNSQYGTIELVL 115
    |||

```

RESULT 4

```

ID MTCP2_MOUSE STANDARD: PRT; 107 AA.
AC Q06945;

```

```

DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE P13 MTCP-1 PROTEIN (MATURE T-CELL PROLIFERATION-1 TYPE B1) (MTCP-1
DE TYPE B1) (P13MTCP1).
GN MTCP1 OR C6.1B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;

```

```

RN 11
RP SEQUENCE FROM N.A.
RT TISSUE-T-cell;
RC MEDLINE=96202516; PubMed=8634440;
RA Madani A., Choukroun V., Soulier J., Cacheux V., Claissé J.-F.,
RA Valensi F., Daliphard S., Cazin B., Lévy V., Leblond V.,
RA Daniel M.-T., Sigaux F., Stern M.-H.;
RT "Expression of p13MTCP1 is restricted to mature T-cell proliferations
RT with C(4;14) translocations."
RL Blood 87:1923-1927(1996).

```

```

CC -1- ALTERNATIVE PRODUCTS: ALTERNATIVE SPLICING PRODUCES TWO SHORT, YET
CC COMPLETELY DIFFERENT PROTEINS. THE LONGER PRODUCT, TYPE-B1, IS
CC KNOWN AS P13 MTCP-1 (AC Q060945). THE SHORTER PRODUCT, TYPE-A, IS
CC TISSUE SPECIFICITY: NOT FOUND AT A SIGNIFICANT LEVEL IN ANY
CC TISSUE.

```


KW Multigene family.
SO SEQUENCE 122 AA; 14170 MW; BD1501F81C24F230 CRC64;

Query Match 25.6%; Score 175.5; DB 1; Length 122;
Best Local Similarity 36.7%; Pred. No. 7,8e-12;
Matches 40; Conservative 12; Mismatches 50; Indels 7; Gaps 1;

OY 12 PPRFLVCTRDITDDEGSRVYVYVRRPNSRRMARASQGSRYEPSTTVHLMQAVHTR 71
DB 13 PPRFLVCTRDITDDEGSRVYVYVRRPNSRRMARASQGSRYEPSTTVHLMQAVHTR 65
OY 72 ELLSSQMFPSQLPAVWOLYPRGKRYRAADSSFWFIADHGQIDSMQVLY 120
DB 66 EPPPOQPINNNSLPTWRLRESNMTYTGDTGYWRLDHSQMCDYQLTL 114

RESULT 7
TCL1_MOUSE STANDARD; PRT; 116 AA.

AC P56280;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE T-CELL LEUKEMIA/LYMPHOMA PROTEIN 1A (P14 TCL1 PROTEIN) (TCL1 ONCOGENE)
DE (TCL-1 PROTEIN).
GN TCL1A OR TCL1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97430049; PubMed=9285687;
RA Narducci M.G., Virgilio L., Engles J.B., Buchberg A.M., Billips L.,
RA Facchiano A., Croce C.M., Russo G., Rothstein J.L.;
RT "The murine Tcl1 oncogene: embryonic and lymphoid cell expression."
RL Oncogene 15:919-926(1997).
CC -1- SIMILARITY: BELONGS TO THE TCL1 FAMILY.
CC -1- SIMILARITY: BELONGS TO THE TCL1 FAMILY.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF031956; AAB87461.1; -;
DR EMBL; Y15376; CAAT5599.1; -;
DR HSSP; P56279; JUSG.
DR MGD; MGI:1097166; Tc11.
DR InterPro; IPR002709; TCL1_MTCPI.
DR Pfam; PF01840; TCL1_MTCPI; 2.
DR ProDom; PD015575; TCL1_MTCPI; 1.
KW Microsome.
SQ SEQUENCE 116 AA; 14112 MW; 46DEDD2F973F389A CRC64;

Query Match 24.7%; Score 169; DB 1; Length 116;
Best Local Similarity 32.8%; Pred. No. 3,6e-11;
Matches 40; Conservative 20; Mismatches 44; Indels 18; Gaps 2;

OY 1 MASASVRLGVP--DGRLLIORPGIYDEEGRTWTVVVRFPNSRRMARASQGSRYEPS 58
DB 1 MATRAARAFETRAHNRRLMIWEKTVYIDERSWLPVYIKSN-----EK 44
OY 59 ITVHLMQAVHTRRELSSQMFPSQLPAVWOLYPRGKRYRAADSSFWFIADHGQIDSMQVLY 118
DB 45 FOVLIRQEDVTILGEMSPSQLPVPELPLMWOLYPRDYSQDSMWQVLIYHIFRVEDM 104
OY 119 VL 120

DB 105 LL 106

RESULT 8
TLB5_MOUSE STANDARD; PRT; 121 AA.
ID TLB5_MOUSE
AC P56845;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TLB5 PROTEIN.
GN TLB5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20056259; PubMed=10588720;
RA Hallas C., Pekarsky Y., Itoyama T., Varnum J., Blich R.,
RA Rothstein J.L., Croce C.M.;
RT "Genomic analysis of human and mouse TCL1 loci reveals a complex of
RT tightly clustered genes."
RL Proc. Natl. Acad. Sci. U.S.A. 96:14416-14423(1999).
CC -1- SIMILARITY: BELONGS TO THE TCL1 FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF195493; AAF12806.1; -;
DR MGD; MGI:1351635; Tc1b5.
DR InterPro; IPR002709; TCL1_MTCPI.
DR ProDom; PD015575; TCL1_MTCPI; 1.
KW Multigene family.
SQ SEQUENCE 121 AA; 13668 MW; 32815CDD629B4297 CRC64;

Query Match 23.9%; Score 163.5; DB 1; Length 121;
Best Local Similarity 33.0%; Pred. No. 1,4e-10;
Matches 37; Conservative 14; Mismatches 44; Indels 17; Gaps 2;

OY 19 QRP-----GIVEDEGRTWTVVVRFPNSRRMARASQGSRYEPSTTVHLMQAV 68
DB 9 GRPLPVLYVSVSLGIYDEHNRVLAUVN-----ETSHSHGRNIEFTCVHLDHMTT 61
OY 69 HTRELSSQMFPSQLPAVWOLYPRGKRYRAADSSFWFIADHGQIDSMQVLY 120
DB 62 LPQEPPOQPINNNSLPTWRLRESNMTYTGDTGYWRLDHSQMCDYQLTL 113

RESULT 9

TLB2_MOUSE STANDARD; PRT; 117 AA.

AC P56841;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE TLB2 PROTEIN.
GN TLB2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20056259; PubMed=10588720;
RA Hallas C., Pekarsky Y., Itoyama T., Varnum J., Blich R.,

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/cor_send_an_email_to_license@isb-sib.ch).

DR EMBL; AF195489; AAF12802.1; -.
DR MGD; MG11351609; TC1B2.
DR InterPro; IPR002709; TC1L_MNCPI.
DR Prodom; PD015575; TC1L_MNCPI; 1.
DR Multigene family.
KW
SQ SEQUENCE 117 AA; 13672 MW; AAB0851165E9D0AA CRC64;

Query Match	23.1%;	Score 158.5;	DB 1;	Length 117;
Best Local Similarity	31.0%;	Pred. No. 4.7e-10;		
Matches 36;	Conservative 17;	Mismatches 52;	Indels 11;	Gaps 2;

QY 10 GVPGRKMIQ-----RPGIYEDEGRWTWTVVVRNPSSRREWARASQGRYEPSITVHLW 64
| | | | | | | | : : : : | : |||||
Db 5 GTPPRLLPQVLSTGPGFYEDEHRLRMVAKL-----ETCSHSPYCNKLETCTVTHLW 56

QY 65 QMAVHTPELLSGOMPFSQLPAWVOLYPGRKYRAADSSFWEIADHGQIDSMEOVL 120
| : : | : : | : : | : : | : : | : : | : : | : :
Db 59 QMTRYPOEPAPYPMNYNLEPMTWLASMNTRYRGTFDAMHWRLLNSHQGDYQIL 114

RESULT	10
TCLA_HUMAN	
ID	
TCLA_HUMAN	
CC	STANDARD;
DS6270.	PRT; 114 AA

NC [F502773](#)
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE T-CELL LEUKEMIA/LYMPHOMA PROTEIN 1A (P14 TCL1 PROTEIN) (TCL1 ONCOGENE (TCL-1 PROTEIN))

GN TCE1A OR TCE1A
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;

SEQUENCE FROM N.A.
MEDLINE=95107991. Pubmed=7809072;
Virgilio L., Narducci M.G., Isobe M., Billips L.G., Cooper M.D.,
Croce C.M., Russo G.;
"Identification of the TGL1 gene involved in T-cell
Proc. Natl. Acad. Sci. U.S.A. 91:12530-12534(1994).
[2]

CHARACTERIZATION, AND LOCALIZATION.
MEDLINE=95079394; Pubmed:7987816;
Fu T.-B., Virgilio L., Narducci M.G., Facchiano A., Russo G.,
Croce C.M.;
"Characterization and localization of the TCL-1 oncogene product.";
Cancer Res. 54:6297-6301(1994).
131

RP X-RAY CRYSTALLOGRAPHY (2.5 ÅNGSTRÖMS).
MEDLINE:98179932. PUBMED:9519406.
HOB P., YANG Y.-S., GUIGNARD L.F., PADILLA A., STERN M.-H.,
JONES L.-M., VAN TILBOURGH H.:
"Crystal structure of PL41CL, an oncogene product involved in T-cell
prolymphocytic leukemia, reveals a novel beta-barrel topology.",
Structure 6:147-155(1998).

-1- SUBCELLULAR LOCATION: MICROSOMAL FRACTION;

CC TISSUE SPECIFICITY: RESTRICTED IN THE T-CELL LINEAGE TO IMMATURE
CC THYMOCYTES AND ACTIVATED PERIPHERAL LYMPHOCYTES. PREFERENTIALLY
CC EXPRESSED EARLY IN T- AND B-LYMPHOCYTE DIFFERENTIATION.
CC DISEASE: ACTIVATED IN CHRONIC T-CELL LEUKEMIA (T-CLL) CARRYING
CC T(14;14)(q11;q32) OR T(7;14)(q35;q23) CHROMOSOME TRANSLOCATIONS OR
CC A T(14;14)(q11;q32) CHROMOSOME INVERSION THAT INVOLVES THE T-CELL
CC RECEPTOR ALPHA/Delta LOCUS.
CC -1 SIMILARITY: BELONGS TO THE TOLL FAMILY.

- TISSUE SPECIFICITY: RESTRICTED IN THE T-CELL LINEAGE TO IMMATURE THYMOCYTES AND ACTIVATED PERIPHERAL LYMPHOCYTES. PREFERENTIALLY EXPRESSED EARLY IN T- AND B-LYMPHOCYTE DIFFERENTIATION.
- DISEASE: ACTIVATED IN CHRONIC T-CELL LEUKEMIAS (T-CLL) CARRYING A T(14;14)(011;032) OR T(7;14)(035;032) CHROMOSOME TRANSLOCATIONS OR A T(14;14)(011;032) CHROMOSOME INVERSION THAT INVOLVES THE T-CELL RECEPTOR ALPHA/Delta LOCUSES.
- SIMILARITY: BELONGS TO THE TOLL FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way distorted, modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

DR EMBL; X82240; CAA57708.1; - .
DR PDB; 1JSG; 18-MAR-98.
DR MIM: 186960; -

DR InterPro; IPR002709; TELL_MTCP1.
DR Pfam: PF01840; TELL_MTCP1.

ProDom; PD015575; TCL1_MTCp1; 1.

AM PROTO-oncogene; Chromosomal translocation; Microsome; 3D-structure
SQ SEQUENCE 114 AA; 13459 MW; 90D55ABC97C36D04 CRC64;

Query Match	21.8%;	Score 149;	DB 1;	Length 114;
Best Local Similarity	29.2%;	Pred. No. 4.6e-09;		
Matches	33;	Conservative	21;	Mismatches 45;
				Indels 14;
				Gaps 1

QY 13 PGRMTQPGIYEDEGRTWTVVFRFNSRREMARASQSRYESITVHLMQAVHTR 72
| | | : | | : : : : : | | : |
Db 15 PRLTAMEKFVYLDEKQHWLPLTIEIKRLQ-----LRLVLRREDVYLGR 60

QY 73 LISSGNPFSOLPAWVQLPGRRKYPADSSFWELIAHGQIDSKDEQVLITYPE 12
:: | : | | : | | | : | | : | | | : | : | : | : | :
Db 61 PMRTPLQGSPILTIMMQLPYDGRYSRSSDSFWRVYHINIDGVEMLLELPD 11

RESULT	11	
VID2_AGRTS		
ID	VID2_AGRTS	STANDARD;
AC	P18592; P06521; P06522;	PRT; 447 AA

DT	01-JAN-1988 (Rel. 06, Created)
DT	01-NOV-1990 (Rel. 16, Last sequence update)
DT	01-AUG-1992 (Rel. 23, Last annotation update)
DE	T-DNA BORDER ENDONUCLEASE VIR2 (EC 3.1.-.-)

Agrobacterium tumefaciens

06 Plasmid pTic58.
0C Bacteria; Proteobacteria; alpha subdivision: Rhizobiaceae group

OC Rhizobiaceae; F
OX NCBI TAXID=358:

	SEQUENCE FROM N.A.
RN	[1]
RP	

RX MEDLINE=90301800;
RA Rogowski P M Po

Zyprian E.M., Steed
"Molecules" are character

the 2003-04 season

Line 26.63-kbp region
Plasmid 23:85-106

[2] PRELIMINARY SEQUEN

MEDLINE=85190558;
Hagiya M., Close T

"Identification of
Aerobacterium tumescens

Proc. Natl. Acad.

AND INTEGRATING
INTO THE PLANT

AND INTEGRATION OF A DEFINED SEGMENT (T-DNA) OF TI PLASMID DNA INTO THE PLANT NUCLEAR GENOME. THE VIR D OPERON ENCODES A SITE-


```

DT 01-JAN-1990 (Rel. 13, Last sequence update)
DE 20-AUG-2001 (Rel. 40, Last annotation update)
DE T-DNA BORDER ENDONUCLEASE VIR2 (EC 3.1.-.-).
GN VIR2.
OS Agrobacterium rhizogenes.
OC Plasmid pRI4b.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=359;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=69039712; PubMed=3185501;
RA Hirayama T., Muranaka T., Ohkawa H., Oka A.;
RT "Organization and characterization of the virD2 genes from
RT Agrobacterium rhizogenes.";
RL Mol. Gen. Genet. 213:229-237(1988).
CC -1- FUNCTION: TUMOR FORMATION BY A.TUMEFACIENS INVOLVES THE TRANSFER
CC AND INTEGRATION OF A DEFINED SEGMENT (T-DNA) OF T1 PLASMID DNA
CC INTO THE PLANT NUCLEAR GENOME. THE VIR2 OPERON ENCODES A SITE-
CC SPECIFIC ENDONUCLEASE THAT CLEAVES AT A UNIQUE SITE WITHIN BOTH
CC 24 BP DIRECT REPEATS FLANKING THE T-DNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X12867; CAA31351.1; -.
DR PIR: S06884; S06884.
KW Hydrolyase; Nuclease; Endonuclease; Crown gall tumor; Plasmid;
KW T-DNA.
SQ SEQUENCE 436 AA; 48371 MW; 932AA365E0966AA6 CRC64;

Query Match
Best Local Similarity 10.9%; Score 74.5; DB 1; Length 436;
Matches 34; Conservative 17; Mismatches 43; Indels 39; Gaps 8;

DY 11 VPP-----GRLVLPQGIYE-----DEGRWTWV--VVVF-----NPSRRMAA-- 47
DY 49 LPDDIHELARSWVQETGYDESQDDEPDEQDELTHIIVSFAGTQTAAYEASREMAAE 108
DY 48 -----RASQSRPESTIVHL-----WQMAVTRRELLSSGOMFESOLPAVWQLYPRKRYR 98
DY 109 MGSQAGGGSYNYLTAHFHIDRPHLHVYVNRRELLGHGMLKIS-----RRHPOLNYDA 162
DY 99 ADSSEWEIA-DHG 110
DY 163 LRINMAETSLRHG 175

RESULT 14
VID2 AGRT6 STANDARD; PRT; 424 AA.
AC P06668;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE T-DNA BORDER ENDONUCLEASE VIR2 (EC 3.1.-.-).
GN VIR2.
OS Agrobacterium tumefaciens.
OC Plasmid pTiA6NC.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=358;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=87028239; PubMed=3021341;
RA Yanofsky M.F., Porter S.G., Young C., Albright L.M., Gordon M.P.,
RA Nester E.W.;

```

```

RT "The virD operon of Agrobacterium tumefaciens encodes a site-specific
RT endonuclease."
RT Cell 47:471-477(1986).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=88032822; PubMed=2822660;
RA Jayaswal R.K., Veluthambi K., Galvin S.B., Slightom J.L.;
RT "Double-stranded cleavage of T-DNA and generation of single-stranded
RT T-DNA molecules in Escherichia coli by a virD-encoded
RT border-specific endonuclease from Agrobacterium tumefaciens.";
RN Bacteriol. 169:5035-5045(1987).
CC -1- FUNCTION: TUMOR FORMATION BY A.TUMEFACIENS INVOLVES THE TRANSFER
CC AND INTEGRATION OF A DEFINED SEGMENT (T-DNA) OF T1 PLASMID DNA
CC INTO THE PLANT NUCLEAR GENOME. THE VIRD OPERON ENCODES A SITE-
CC SPECIFIC ENDONUCLEASE THAT CLEAVES AT A UNIQUE SITE WITHIN BOTH
CC 24 BP DIRECT REPEATS FLANKING THE T-DNA.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AF242881; AAA98390.1; -.
DR PIR: M17989; AAA22114.1; -.
DR PIR: B29826; B29826.
DR PIR: B25063; B25063.
KW Hydrolyase; Nuclease; Endonuclease; Crown gall tumor; Plasmid;
KW T-DNA.
SQ SEQUENCE 424 AA; 47546 MW; B2CB7B82C01A6C1C CRC64;

Query Match
Best Local Similarity 10.6%; Score 72.5; DB 1; Length 424;
Matches 38; Conservative 24; Mismatches 49; Indels 47; Gaps 10;

DY 11 VPPGL-----WLPQGIYE-----DEGRWTWV--VVVF-----NPSRRMAA 49
DY 49 VPPDRIELASWVTEGTGYDESQSDDDRDQDELTHIIVSFAGTQTAAYEASREMAAE 108
DY 50 SGG-----RYESITIVHL-----WQMAVTRRELLSSGOMFESOLPAVWQLYPRKRYR 97
DY 109 MGSQAGGGSYNYLTAHFHIDRPHLHVYVNRRELLGHGMLKISRRP--QLWYDGLRK 166
DY 98 AADSEFEIADHGOI-----DSMDVLVTYQPERK 127
DY 167 MAEISL-----RHGIVDANISRAERGIARPTTYAEHRH 200

RESULT 15
GLGB SOLTU STANDARD; PRT; 861 AA.
AC P30924;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 20-AUG-2001 (Rel. 40, Last annotation update)
DE 1,4-ALPHA-GLUCAN BRANCHING ENZYME (EC 2.4.1.18) (STARCH BRANCHING
DE ENZYME) (O-ENZYMES).
GN SBE1 OR SBE.
OS Solanum tuberosum (Potato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
OX NCBI_TaxID=4113;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. DIANEELA;
RX MEDLINE=94105324; PubMed=8278528;
RA Poulsen P., Kreiberg J.D.;
RT "Starch branching enzyme cDNA from Solanum tuberosum.";
RT Plant Physiol. 102:1053-1054(1993).

```

```

RN [2]
RP SEQUENCE OF 279-527 FROM N.A.
RC STRAIN-CV, DESIREE; TISSUE-Tuber;
RX MEDLINE=92079917; Pubmed=1745241;
RA Kossman J., Visser R.G.F., Mueller-Roeber B., Willmitzer L.,
RA Sonnewald U.; expression analysis of a potato cDNA that encodes
RT branching enzyme: evidence for co-expression of starch biosynthetic
RT genes."
RL Mol. Gen. Genet. 230:39-44(1991).
CC -1- CATALYTIC ACTIVITY: FORMATION OF 1,6-GLUCOSIDIC LINKAGES OF
CC STARCH.
CC -1- PATHWAY: THIRD STEP IN STARCH BIOSYNTHESIS.
CC -1- SUBUNIT: MONOMER.
CC -1- SUBCELLULAR LOCATION: AMYLOPLAST.
CC -1- SIMILARITY: BELONGS TO FAMILY 13 OF GLYCOSYL HYDROLASES, ALSO
CC KNOWN AS THE ALPHA-AMYLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X69805; CAA49463.1; -
CC PIR; S18594; S18594.
CC DR Mendel; 13376; SOLtu;Sbel;1.
CC DR InterPro; IPR000461; Alpha-amylase.
CC Pfam; PF00128; alpha-amylase; 1.
CC Starch biosynthesis; Transferase; Glycosyltransferase; Amyloplast.
CC ACT SITE 424 424 BY SIMILARITY.
CC FT ACT SITE 484 484 BY SIMILARITY.
CC FT ACT SITE 553 553 BY SIMILARITY.
CC SEQIDENCE 861 AA; 99083 MW; F3D519AC7CF1BEF2 CRC64;
SO

```

```

Query Match 10.28; Score 70; DB 1; Length 861;
Best Local Similarity 24.2%; Pred. No. 11;
Matches 31; Conservative 17; Mismatches 34; Indels 46; Gaps 8;

```

```

OY 17 WIORPGIYDEGRTWTVVVRFPNSRREMAR-SQGSRYEPTIVHLQMAVHTRE---72
DB 620 WIDFP-----RCNNW-----SYDKCRQWNLADSEHLRYK--FWNAFDRAMNSLDEKF 666
OY 73 -LISGQMPFS-----QLPAYQIYGR-----KYRAA-DSEF 103
DB 667 SFLASGKQIVSSMDDNKVVVERGDLVVFNFHKNITGCIYVCGDLPKRYVALDSDA 726
OY 104 WEIADHQ 111
DB 727 WEFGGHR 734

```

Search completed: November 29, 2001, 04:06:42
Job time: 208 sec

Query Match	98.88;	Score 677.	DB 4;	Length 126;
Best Local Similarity	99.3%;	Pred. No. 1.3e-62;		
Matches 127;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
Qy	1	MASASVRLGVPGRGLTQRPGLTYDEEGCPWTVYVVRNPSREKMARASQCSREPST	60	
Db	1	MASASVRLGVPGRGLTQRPGLTYDEEGCTWTVYVVRNPSREKMARASQCSREPST	60	
Oy	61	VHLMQAVHTRELLSSGQCFPSQLPAVWQLPGPKTRADSSFEINIDGQIDSMEDVYL	120	
Db	61	VHLMQAVHTRELLSSGQCFPSQLPAVWQLPPRKTRADSSFEINIDGQIDSMEDVYL	120	

```

Oy      121 TYOPERKD 128
        |||||
Db      121 TYOPERKD 128

RESULT          2
O90XN9         PRELIMINARY; PRT: 107 AA.
AC O90XN9;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
Dt 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
Dt 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
Df T-CELL LEUKEMIA PROTEIN TCL1B3A.
Gn TCL1B3.
Os Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Ox NCBI_Taxid=10090;
Rn [1]
Rp SEQUENCE FROM N.A.
Ra MEDLINE=20056259; PubMed=10588720;
Rd Hallas C., Pekarsky Y., Itoyama T., Varum J., Bichi R.,
Ra Rothenstein J.L., Croce C.M.;
Pt "Genomic analysis of human and mouse TCL1 loci reveals a complex of
rt tightly clustered genes.";
Rl Proc. Natl.Acad. Sci. U.S.A. 96:14418-14423(1999).
DR MGI:AF195491; AAI12804.1; -.
Dr MGI: MGI1351600:TCL1B3.
Sq SEQUENCE 107 AA; 12224 MW; DDC2AMC489D02DBE CRC64;

Query Match           16.5%; Score 113; DB 11; Length 107;
Best Local Similarity 31.6%; Pred. No. 0.00022;
Matches 30; Conservative 13; Mismatches 46; Indels 6; Gaps

Oy      29 GRWVTVVFNFNSREWRAS--QGSRYSPTIVHMQAVHTRELLSSGMPFSQLP 85
        | : : : ||| | | | | | | | | | | | | | | | | | | | | | | |
Db      8 GLGNAS-LFES-GRRALGRASLKLSCHLEFCITVHLQHHTTPEPPPOOPINNSLP 64
        :::: | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Oy      86 AVMDLYPGKRYRADDSFWELADHGDISMEDLVYL 120
        :::: | : : : | : : : | : : : | : : : | : : : | : : : | : : :
Db      65 TWMRLESMTYTGTDGTIWLRLDHSDNGDTLDLIL 99

RESULT          3
O9F587         PRELIMINARY; PRT: 448 AA.
AC O9F587;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
Dt 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
Dt 01-MAR-2001 (TREMBLrel. 16, Last annotation update)
Df RIOFP168 PROTEIN.
Gn RIOFP168.
Os Agrobacterium rhizogenes.
OC Plasmid pRI1724.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
Ox NCBI_TaxID=359;
Rn [1]
Rp SEQUENCE FROM N.A.
Ra STRAIN=MAFF03-01724;
Rd Moriyuchi K., Meeda Y., Satou M., Satuli N., Kataoka M., Tanaka N.,
Ra Toshida K.;
Pt "The complete nucleotide sequence of a Ri (root inducing) plasmid
rt indicates its chimerical structure between Ti and Sym plasmids.";
Rt Submitted (MAY-2000) to the EMBL/GenBank/DDBB databases.
Rn [2]
Rp SEQUENCE FROM N.A.
Ra STRAIN=MAFF03-01724;
Rd Moriyuchi K., Meeda Y., Satou M., Kataoka M., Tanaka N., Yoshida K.;
Rt "Analysis of unique variable region of a plant root inducing plasmid,
rt pRI1724, by the construction of its physical map and library.";
Rt Submitted (MAY-2000) to the EMBL/GenBank/DDBB databases.
```

RN SEQUENCE FROM N.A.
 RP [3]
 RA STRAIN=MAFF03-01724;
 RC Moriyuchi K., Nishida T., Maeda Y., Tanaka N., Yoshida K.;
 RD "Genome structure of RI plasmid (1): Construction of linking library
 RE and physical map of pRI1724 in Japanese *Agrobacterium*.";
 RF Nucleic Acids Symp. Ser. 39:189-190(1998).
 RL [4]
 RN SEQUENCE FROM N.A.
 RP STRAIN=MAFF03-01724;
 RC MEDLINE=20241294; PubMed=10780382;
 RA Maeda Y., Moriyuchi K., Kataoka M., Satou M., Satuli N., Tanaka N.,
 RD Yoshida K.;
 RE "Genome structure of RI plasmid (1): Sequencing analysis of T-DNA and
 RF its flanking regions of pRI1724 in Japanese *Agrobacterium*
 RG rhizogenes.";
 RH Nucleic Acids Symp. Ser. 42:67-68(1999).
 RL EMBL=AP002086; BAB16287.1; -.
 RN Plasmid.
 KW
 SQ SEQUENCE 448 AA; 50021 MW; FB2594921BCB704 CRC64;

	Query Match	11.8%;	Score 81;	DB 2;	Length 448;
	Best Local Similarity	26.8%;	Pred No. 2.4;		
	Matches 40; Conservative	16;	Mismatches 49;	Indels 44;	Gaps 9.
Oy	11 VPP-----GRLMIORPGIYE---DEEGRTWVT--VVVRF-----NPSRRRMAA-- 47 : : : : : : : : : : : : Db	49 VPPEIRELARSWLQETGYHESOPDEREQDLTHITIVSFPAIGTSOVAAYASREMAAE 108 			
Oy	48 ----RASGSRYEPSITSYHL-----WMAMVHTRELSSGMFSPQLPVMOLYPGRKYRA 98 : : : : : : : : : : : : : : Db	109 MESSAGGGCRNYNTAFHIDNRPHLLVYNVRRELLGHGWMLKIS-----RBHPOLNYDA 162 			
Oy	99 ADSEFWELA-DHG 116 : : : : : : : : : : : : Db	163 LRIMKEISLRHGVS-----LEATFRKER 186 			
RESULT	4				
ID	044461	PRELIMINARY:	PRZ,	447 AA.	
AC	044461;				
DJ	01-NOV-1996 (TREMBlrel. 01. Created)				
DT	01-NOV-1996 (TREMBlrel. 01. Last sequence update)				
DD	01-MAR-2001 (TREMBlrel. 16, Last annotation update)				
DE	UNKNOWN PROTEIN Z.				
OC	Aerobacterium radiobacter.				
OS	Bacteria, Proteobacteria; alpha subdivision; Rhizobiaceae group;				
OX	Rhizobiaceae; Rhizobium.				
NX	NCBI_TaxID=358;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RA	Wang K., Herrera-Estrrella A.H.; Van Montagu M.M.;				
RL	Submitted (JUL-1990) to the EMBL/Genbank/DBJ databases.				
DR	EMBL; M33673; AAA2111.1;				
SQ	SEQUENCE 447 AA: 49649 MW; 9P43JA501B4Z2S5 CRC64;				
	Query Match	11.6%;	Score 79.5;	DB 2;	Length 447;
	Best Local Similarity	26.3%;	Pred. No. 3.4;		
	Matches 35; Conservative	16;	Mismatches 43;	Indels 39;	Gaps 6
Oy	11 VPP-----GLMIDIRPGITE----DECGRWWT--VVVVF-----NPSRRRMAA-- 47 : : : : : : : : : : : : Db	49 LPPOIHELLARMWGERTGESQPDEREQQDELTHITIVSFPAIGTSOVAAVAASREMAAE 108 			
Oy	48 ---RASGSRYEPSITSYHL-----WMAMVHTRELSSGMFSPQLPVMOLYPGRKYRA 98 : : : : : : : : : : : : : : Db	109 MGSGAGGAGRYYNTAFHIDRDPHLLVYNVRRELLGHGWMLKIS-----RBHPOLNYDA 162 			
Oy	99 ADSEFWELA-DHG 110 : : : : : : : : : : : : Db				

DB 163 LRKMAEISLRHG 175

RESULT 5

ID Q9R6B9 PRELIMINARY; PRT; 447 AA.

AC Q9R6B9

DT 01-MAY-2000 (TREMBLrel. 13, Created)

DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)

DE TIORF156 PROTEIN.

GN TIORF156

OS Agrobacterium radiobacter.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Rhizobium.

OX NCBI_TaxID=358;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-MAF301001;

RX MEDLINE=20184752; PubMed=10721727;

RA Suzuki K., Hattori Y., Uraji M., Ohta N., Iwata K., Murata K.,

Kato A., Yoshida K.;

RT "Complete nucleotide sequence of a plant tumor-inducing Ti plasmid.";

RL Gene 242:331-336(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-MAF301001;

RX MEDLINE=98193120; PubMed=9524202;

RA Suzuki K., Ohta N., Hattori Y., Uraji M., Kato A., Yoshida K.;

RT "Novel structural difference between nopaline- and octopine- type trb1

gene: construction of genetic and physical map and sequencing of

trb1/trai and rep gene clusters of a new Ti plasmid pTi-SAKURA.";

RL Biochim. Biophys. Acta 1396:1-7(1998).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-MAF301001;

RA Hattori Y., Suzuki K., Ohta N., Uraji M., Kato A., Yoshida K.;

RT "Genome structure of pTi-SAKURA (11): Strategy for DNA sequencing of a

Japanese cherry-Ti plasmid.";

RL Nucleic Acids Symp. Ser. 37:159-160(1998).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN-MAF301001;

RA Ohta N., Suzuki K., Hattori Y., Uraji M., Kato A., Yoshida K.;

RT "Genome structure of pTi-SAKURA (11): Characteristics of T-DNA.";

RL Nucleic Acids Symp. Ser. 39:185-186(1998).

RN [5]

RP SEQUENCE FROM N.A.

RC STRAIN-MAF301001;

RA Uraji M., Suzuki K., Ohta N., Hattori Y., Kato A., Yoshida K.;

RT "Genome structure of pTi-SAKURA (1V): Characteristics of tra region.";

RL Nucleic Acids Symp. Ser. 39:187-188(1998).

RN [6]

RP SEQUENCE FROM N.A.

RC STRAIN-MAF301001;

RA Hattori Y., Suzuki K., Ohta N., Uraji M., Kato A., Yoshida K.;

RT "Genome structure of pTi-SAKURA (V): Complete nucleotide sequence of

Ti plasmid pTi-SAKURA's vir region in Agrobacterium tumefaciens.";

RL Nucleic Acids Symp. Ser. 39:265-266(1998).

DR EMBL: AB016260; BAA87781.1; -.

KW Plasmid.

SQ SEQUENCE 447 AA; 43603 MW; 5A9E646F2D8894AC CRC64;

Query Match 11.6%; Score 79.5; DB 2; Length 447;

Best Local Similarity 26.3%; Pred. No. 3.4;

Matches 35; Conservative 16; Mismatches 43; Indels 39; Gaps 8;

DB 49 LPPQIHLANSWQETGYDESQPDREKQELTTHIVSFAGTSQVAAVAAASREMAAE 108

QY 48 ---RASQGSRYESTVHL-----NOMAVHRELLSSGMPSPQPAVMOLYGRKRYA 98

DB 109 MFGSAGGGRYNTLAFHIDRPHLHVYVNRRELLGHGWLKIS-----RRHPQLVNDA 162

QY 99 ADSSFWIEA-DHG 110

DB 163 LRKMAEISLRHG 175

RESULT 6

ID Q9HYZ1 PRELIMINARY; PRT; 238 AA.

AC Q9HYZ1

DT 01-MAR-2001 (TREMBLrel. 16, Created)

DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)

DE PROBABLE TRANSCRIPTIONAL REGULATOR.

GN PA3249.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

OX NCBI_TaxID=287;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-PA01;

RX MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,

Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

Garber R.L., Goltz S.N., Tolentino E., Westbrook-Wadman S., Yuan Y.,

Brody L.L., Coulter S.N., Folger K.R., Kas A., Lathig K., Lim R.M.,

Smith K.A., Spencer D.H., Mong G.K.-S., Wu Z., Paulsen I.T.,

Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

opportunistic pathogen.";

RL Nature 406:959-964(2000).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN-PA01;

RX MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,

Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

Garber R.L., Goltz S.N., Tolentino E., Westbrook-Wadman S., Yuan Y.,

Brody L.L., Coulter S.N., Folger K.R., Kas A., Lathig K., Lim R.M.,

Smith K.A., Spencer D.H., Mong G.K.-S., Wu Z., Paulsen I.T.,

Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

opportunistic pathogen.";

RL Nature 406:959-964(2000).

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN-PA01;

RX MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,

Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

Garber R.L., Goltz S.N., Tolentino E., Westbrook-Wadman S., Yuan Y.,

Brody L.L., Coulter S.N., Folger K.R., Kas A., Lathig K., Lim R.M.,

Smith K.A., Spencer D.H., Mong G.K.-S., Wu Z., Paulsen I.T.,

Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;

RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an

opportunistic pathogen.";

RL Nature 406:959-964(2000).

RN [4]

Query Match 11.5%; Score 79; DB 2; Length 238;

Best Local Similarity 23.9%; Pred. No. 1.8;

Matches 28; Conservative 17; Mismatches 46; Indels 26; Gaps 3;

QY 5 ASYRLGVPPRLMIORQIYDEGRTWTVVFRNPSRRRARRASQGSRYESTVHL 64

DB 42 ATRTILREALIOLISOGILYREE-----RKGWFSERPLAVNPLVNSHF 87

QY 65 OMA-----VHTRELLSSGMPF-----OLPAVMOLYGRKRYARRASSWELACH 109

DB 88 AMVSEGRVPRATEVLSARQNPASANCLELPLSSVIOIRARVVGRLVLYEH 144

RESULT 7

ID 070681 PRELIMINARY; PRT; 588 AA.

AC 070681

DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)

DE COAT PROTEIN (FRAGMENT).

GN Sugarcanne streak mosaic virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Potyviridae;

OC Potyvirus.

OX NCBI_TaxID=53954;

RN [1]

NX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-BRAIN;
 RX MEDLINE=20039619; PubMed=10574462;
 RA Nagase T., Ishikawa K., Kikuno R., Hirose M., Nomura N., Ohara O.;
 RT "Prediction of the coding sequences of unidentified human genes. XV.
 The complete sequences of 100 new cDNA clones from brain which code
 for large proteins in vitro."
 RL DNA Res. 6:337-343(1999).
 DR EMBL; AB033096; BAA86584.1;
 DR InterPro: IPR002106; AA:trna_ligase_II.
 DR InterPro: IPR002318; trna-synt_2c.
 DR Pfam: PF01411; trna-synt_2c; 1.
 DR PRINTS: PR00980; TRNASYNTHALA.
 DR PROSITE: PS00339; AA:trna_ligase_II_2; UNKNOWN_1.
 FT NON_TER 1
 SQ SEQUENCE 986 AA; 107427 MW; E6FDD207907A831 CRC64;

Query Match 10.9%; Score 75; DB 4; Length 986;
 Best Local Similarity 26.0%; Pred. No. 25;
 Matches 34; Conservative 16; Mismatches 51; Indels 30; Gaps 8;

QY 10 GVPGRMTQ-----RGTEDEDEGR-TWTVY-----VYRNPSPRRMARASGSKRE 56
 DB 162 GIPERLMTISYEDGPKAGLPDLETRDLSLGPASRVLSFGPOENWEGDTGP-CG 220
 QY 57 PSIVHLMQAVHTELLSSGMPFSQLPAYWOLYPRKRYRAADSPWEI----ADHGOI 112
 DB 221 PCTEIH-YDLA-----GGVGPQLVELMNLVFMQHNEDAGSLQPLPQRHVDTCM- 269
 QY 113 DSMEOVLVTVQ 123
 DB 270 -GLERLVAVLQ 279

RESULT 11
 Q9HXM7 PRELIMINARY; PRT; 463 AA.
 AC Q9HXM7
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE PROBABLE METALLO-OXIDOREDUCTASE.
 GN PA3768.
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 NX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=PAO1;
 RX MEDLINE=20437337; PubMed=10984043;
 RA Stever C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 Gader R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 Reiter J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PAO1, an
 opportunistic pathogen."
 RL Nature 406:959-964(2000).
 DR EMBL; AE004795; AAG07155.1;
 DR InterPro: IPR001117; Cu-oxidase.
 DR InterPro: IPR002355; Multicu_oxidase2.
 DR Pfam: PF00394; Cu-oxidase; 2.
 DR PROSITE: PS00079; MULTICOPPER_OXIDASE; 1.
 DR PROSITE: PS00080; MULTICOPPER_OXIDASE2; 1.
 KW Complete proteome.
 SQ SEQUENCE 463 AA; 51448 MW; 44222B5DF9EBB3 CRC64;

Query Match 10.7%; Score 73; DB 2; Length 463;
 Best Local Similarity 24.3%; Pred. No. 17;
 Matches 35; Conservative 18; Mismatches 55; Indels 36; Gaps 8;

QY 1 MASASVRLGVPGRMTQPRGI-----YEDEGRTWTVY-----VRF 39
 DB 12 LAGLAAYGAGAGATMLNARQVAQETDELTAPELTPGFSPPALATGGCPVEL 71
 QY 40 NPSRRMARASGSKREPSITVHLMQAVHTELLSSGMPFSQLPAYWOLYPR- -K 95
 DB 72 RAKGEWLRLRFRTNRLDEPTIH-WH---GIRLPIEMDGVYISOPPV---QPGSEFTYQ 124
 QY 96 YRAADS-SFWEIADHGOIDSMEO 118
 DB 125 FKTDAGSTWY---HPHLSSEQL 145

RESULT 12
 Q9FT99 PRELIMINARY; PRT; 464 AA.
 AC Q9FT99
 DT 01-MAR-2001 (TREMblrel. 16, Created)
 DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE GLUCURONOSYL TRANSFERASE-LIKE PROTEIN.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eucots II; Brassicales; Brassicaceae; Arabidopsis.
 NX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=COLUMBIA;
 RX MEDLINE=9397451; PubMed=10470850;
 RA Kaneo T., Katoh T., Sato S., Nakamura Y., Asamizu E., Kotani H.,
 Miyajima N., Tabata S.;
 RT "Structural analysis of Arabidopsis thaliana chromosome 5. IX.
 Sequence features of the regions of 1,011,550 bp covered by seventeen
 P1 and TAC clones."
 RL DNA Res. 6:183-195(1999).
 DR EMBL; AB017060; BAB10792.1; -
 DR InterPro: IPR002213; UDPGT.
 DR Pfam; PF00201; UDPGT; 1.
 KW transferase.
 SQ SEQUENCE 464 AA; 52360 MW; 522E2121P643896D CRC64;

Query Match 10.7%; Score 73; DB 10; Length 464;
 Best Local Similarity 23.8%; Pred. No. 17;
 Matches 31; Conservative 18; Mismatches 49; Indels 32; Gaps 8;

QY 16 LMIORPGIYDEDEGRTWY-----VYRNPSPRRMARASGSKREPSITV 61
 DB 303 LMVVRPG---SVHGDMIESLPQSFESIDGKGRIVRWAPQLDVLAHRTAG---PLTH 355
 QY 62 HLMQAVHTELLSSGMPFSQLPAYWOLYPRKRYRAADSPFEIDH--GOID--SMEQ 117
 DB 336 NGMN---SILESICG-VPMICLPCKMDQFVNAR--ISEVVRVGHLEGRTERTER 408
 QY 118 LVTVQPERK 127
 DB 409 AVIRLWESK 418

RESULT 13
 Q9L803 PRELIMINARY; PRT; 513 AA.
 AC Q9L803
 DT 01-OCT-2000 (TREMblrel. 15, Created)
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)
 DE PUTATIVE ACYL-COA DEHYDROGENASE.
 GN PDRFO.

OS Pseudomonas stutzeri (Pseudomonas perfectomarina).
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
 OC Pseudomonas.
 OX NCBI_TaxID=316;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=KC;
 RA Lewis T.A., Cortese M.S., Sebat J.L., Green T.L., Crawford R.L.;
 RT "Identification of a Region of the Pseudomonas stutzeri strain KC
 Chromosome Containing Genes for the Biosynthesis of Pyridine-2,6-
 bis(thiocarboxylic acid), the Agent of Carbon Tetrachloride
 RT Dechlorination Produced by this Organism."
 RT Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AF196567; AAF3139.1; -
 DR InterPro: IPR001552; ACYL-COA_dh.
 DR Pfam: PF00441; ACYL_COA_dh_1.
 DR PROSITE: PS00072; ACYL_COA_dh_1; UNKNOWN_1.
 DR PROSITE: PS00073; ACYL_COA_dh_2; 1.
 SQ SEQUENCE 513 AA; 57432 MW; 4F6D6E16F3D5C34 CRC64;

Query Match 10.7%; Score 73; DB 2; Length 513;
 Best Local Similarity 24.3%; Pred. No. 19;
 Matches 25; Conservative 16; Mismatches 42; Indels 20; Gaps 4;
 QY 39 FNPSPREMARASQGR-----YESP-----TVHMQMAVHTRELLSSGOMP 80
 DB 45 FGPRTSWMSKSGKAGRLVHCHHPQATCARRRNRQVVSCHLHLHHPKRRRANGVP 104
 QY 81 FSOLPAWOLYR-GRKYRAADS-SFWEIADHGQIDMSOVLVT 121
 DB 105 LGQLPRKQOLDPSGRSLSLACTPANDMDISLPNEVQMLVST 147

RESULT 14
 Q9EMP5 PRELIMINARY; PRT; 2229 AA.
 AC Q9EMP5;
 DT 01-MAR-2001 (TREMBLrel. 16, Created)
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PUTATIVE NON-RIBOSOMAL PEPTIDE SYNTHASE.
 GN SC4C2.17.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Brown S.P., Harris D.;
 RT Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
 RT Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kleser H.M., Denaplatte D., Eichner A., Cullum J.,
 RA Kinashl H., Hopwood D.A.;
 RT "A set of ordered cosmid and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 DR EMBL: AL450432; CA017499.1;
 DR InterPro: IPR000873; AMP-bind.
 DR InterPro: IPR001242; DUF4.
 DR InterPro: IPR003880; Phosphopantl_attach.
 DR Pfam: PF00501; AMP-binding_2.
 DR Pfam: PF00668; Condensation; 3.
 DR Pfam: PF00550; pp-binding; 3.

DR PROSITE: PS00075; ACP_DOMAIN; 3.
 DR PROSITE: PS00012; PHOSPHOPANTETHEINE; UNKNOWN_1.
 KW Phosphopantetheine.
 SQ SEQUENCE 2229 AA; 239295 MW; 66EB782247C326CD CRC64;

Query Match 10.7%; Score 73; DB 2; Length 2229;
 Best Local Similarity 28.4%; Pred. No. 1e+02;
 Matches 25; Conservative 6; Mismatches 29; Indels 28; Gaps 4;
 QY 6 SVRLGVP-----PERLMIORPGI---YDEEGRTWTVVVR-NP 41
 DB 878 SVYGVPLNMRNARVYDGRGNDPDLPGELMIGSPGVANGYGDGR---AERVVDH 933
 QY 42 SREMARASQGRYEPSTVHMQMAVH 69
 DB 934 DGERWYRSGDLARYPDGVLEFLGRADH 961

RESULT 15
 O69839
 ID O69839 PRELIMINARY; PRT; 513 AA.
 AC O69839;
 DT 01-AUG-1998 (TREMBLrel. 07, Created)
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE PUTATIVE IRON-SULFUR BINDING OXIDOREDUCTASE.
 GN SCIB5.11C.
 OS Streptomyces coelicolor.
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OX NCBI_TaxID=1902;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Harris D., Taylor K.;
 RT Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RA Parkhill J., Barrell B.G., Rajandream M.A.;
 RT Submitted (MAY-1998) to the EMBL/Genbank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=A3(2);
 RX MEDLINE=97000351; PubMed=8843436;
 RA Redenbach M., Kleser H.M., Denaplatte D., Eichner A., Cullum J.,
 RA Kinashl H., Hopwood D.A.;
 RT "A set of ordered cosmid and a detailed genetic and physical map for
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."
 DR EMBL: AL023517; CA018985.1; -
 DR HSSP: P08980; IRES.
 DR InterPro: IPR00205; NAD_binding.
 DR InterPro: IPR001281; Rieske.
 DR Pfam: PF00355; Rieske_1.
 DR PRINTS: PR00162; RIESKE.
 DR PROSITE: PS00199; RIESKE_1; 1.
 DR PROSITE: PS00200; RIESKE_2; 1.
 SQ SEQUENCE 513 AA; 34931 MW; 06F0967279C9FAF CRC64;

Query Match 10.6%; Score 72.5; DB 2; Length 513;
 Best Local Similarity 32.8%; Pred. No. 21;
 Matches 19; Conservative 7; Mismatches 19; Indels 13; Gaps 2;

QY 11 VPRGRLMIORPG-----YDEEGR-----TWYTVVYRFPSPREMARASQGRSX 55
 DB 436 LPPEGAVVRAAGGRLAVYRDEGALHAVSPRCHLGLVDFAERAERACPCGSRF 493

Fri Nov 30 11:03:13 2001

Job time: 222 sec

us-09-526-329-39.rspt

